Risk-Averse Stochastic Programming vs. Adaptive Robust Optimization: A Virtual Power Plant Application

Ricardo Lima, Antonio Conejo, Loïc Giraldi, Olivier Le Maitre, Ibrahim Hoteit, Omar Knio

Risk-averse stochastic programming vs. adaptive robust optimization: a virtual power plant application

Ricardo M. Lima[a,*], Antonio J. Conejo[b], Loïc Giraldi[c], Olivier Le Maître[d], Ibrahim Hoteit[e], Omar M. Knio[a]

[a] *Computer, Electrical and Mathematical Sciences & Engineering Division, King Abdullah University of Science and Technology, (KAUST), Thuwal 23955-6900, Saudi Arabia*
[b] *Integrated Systems Engineering-Electrical and Computer Engineering, The Ohio State University, OH, USA*
[c] *CEA, DES, IRESNE, DEC, Cadarache F-13108 Saint-Paul-Lez-Durance, France*
[d] *Centre de Mathématiques Appliquées, CNRS, Inria, Ecole Polytechnique, Palaiseau, France*
[e] *Physical Science and Engineering Division, King Abdullah University of Science and Technology, (KAUST), Thuwal 23955-6900, Saudi Arabia*

Abstract

This paper compares risk-averse optimization methods to address the self-scheduling and market involvement of a virtual power plant (VPP). The decision-making problem of the VPP involves uncertainty in the wind speed and electricity price forecast. We focus on two methods: risk-averse two-stage stochastic programming (SP) and two-stage adaptive robust optimization (ARO). We investigate both methods concerning formulations, uncertainty and risk, decomposition algorithms, and their computational performance. To quantify the risk in SP, we use the conditional value at risk (CVaR) because it can resemble a worst-case measure, which naturally links to ARO. We use two efficient implementations of the decomposition algorithms for SP and ARO, and we assess 1) the operational results regarding first-stage decision variables, estimate of expected profit, and estimate of the CVaR of the profit; and 2) their performance taking into consideration different sample sizes and risk management parameters. The results show that similar first-stage solutions are obtained depending on the risk parameterizations used in each formulation. Computationally, we identified three cases: 1) SP with a sample of 500 elements is competitive with ARO; 2) SP performance degrades comparing to the first case and ARO fails to converge in four out of five risk parameters; 3) SP fails to converge, whereas ARO converges in three out of five risk parameters. Overall, these performance cases depend on the combined effect of deterministic and uncertain data, and risk parameters.

Keywords: Stochastic programming, robust optimization, risk management, virtual power plant

1. Introduction

This work lies at the intersection of two subjects that have been receiving increasing attention in the literature: the optimal operation of virtual power plants (VPPs) and optimization under uncertainty. One of the first works to discuss the concept of VPP was the book edited by Awerbuch and Preston [1], where several topics related with VPPs were addressed. Later, VPPs were discussed in Pudjianto et al. [2], where a VPP is defined as an aggregation of distributed energy resources that aggregates capacity and creates a single entity to participate in the electricity market, as well as to support transmission system management.

The considered VPP consists of a thermal plant, a pump-storage hydro plant, and a wind farm that interacts with the electricity market as a single entity. The VPP aims to maximize its profit and faces

*Corresponding author
Email addresses: ricardo.lima@kaust.edu.sa (Ricardo M. Lima), conejonavarro.1@osu.edu (Antonio J. Conejo), loic.giraldi@cea.fr (Loïc Giraldi), olivier.le-maitre@polytechnique.edu (Olivier Le Maître), ibrahim.hoteit@kaust.edu.sa (Ibrahim Hoteit), omar.knio@kaust.edu.sa (Omar M. Knio)

Article accepted to INFORMS Journal on Computing (2022)

two interrelated problems: 1) the self-scheduling of the pump-storage hydro and thermal plants; and 2) the interaction of the VPP with the electricity market. The self-scheduling problem determines when the pump-storage hydro plant generates or consumes power and the thermal plant power output for each period. The electricity market interaction involves selecting forward contracts and the sale or purchase of electricity in the hourly electricity pool. The detailed VPP problem statement and the underlying assumptions are provided in Section 2.

Conceptually, the main advantages of VPPs are the following: a) mitigation of risk of low or negative profits for small generators (when aggregated into a VPP) trading in the electricity market; b) increased overall capacity of the aggregated generators, compared with the individual units, which leads to economies of scale and allows small generators to participate in electricity markets with minimum capacity requirements; c) and an increase of efficiency of the overall power system; see Pudjianto et al. [2] for a discussion on objectives and advantages of the concept of VPP. More specifically, Dietrich et al. [3] evaluated the impact of VPPs in the Spanish power system and concluded that the VPPs did not have a relevant impact on the generators of the system that were not part of VPPs. However, the aggregation into VPPs had a significant impact on the generators that were aggregated, in terms of their benefits maximization.

The growing interest on VPPs is also motivated by the flexibility of the aggregation of renewable energy sources (RES) and conventional sources and their complementarity [4]. The presence of RES in the composition of VPPs and the volatility of the electricity prices in the electricity market make uncertainty to play a critical role in decision-making for VPPs.

Optimization problems with uncertain parameters have received increasing attention, mainly stochastic programing (SP) formulations involving risk measures [5, 6, 7] and robust optimization [8, 9, 10, 11, 12].

In the next two sub-sections, we discuss some works that have proposed optimization models for VPP problems.

1.1. Stochastic programming formulations for VPPs

Pandžić et al. [13] was one of the first works to propose a SP formulation for an offering model for a VPP (aggregating a wind power plant, a quick-response conventional power plant, and a pump-storage hydro plant). They developed a two-stage SP formulation that maximizes the expected profit of a VPP, and determines optimal offers to the day-ahead market and balancing market. Their solution approach is based on the solution of the extensive formulation of the SP problem, and a risk measure was not used in their formulation. Tajeddini et al. [14] have also considered the optimal operation of a VPP (aggregating a wind power plant, a photovoltaic system, a micro turbine, a diesel generator, and a battery bank), which participates in the day-ahead and balancing markets. They proposed a two-stage risk-averse SP formulation using the conditional value at risk (CVaR) in the objective function to avoid the risk of low profits. The resulting Mixed-Integer Linear Programming (MILP) problem was solved using an extensive formulation and an MILP solver. Riveros et al. [15] addressed the optimal bidding strategies for a VPP (aggregating combined heat and power based system for district heating and renewable energy generators) using a two-stage risk-neutral SP formulation. They considered that the VPP bids in the day-ahead and balancing markets, and compared alternative bidding strategies that differ on the re-scheduling and adaptiveness of the bidding strategy as uncertainty is revealed. Kardakos et al. [16] studied a different offering problem for a VPP (aggregating a wind power farm, consumers, and a battery storage system), which considered strategic offering decisions to influence the prices and maximize a combination of expected profit and CVaR of the profit. Their problem is modeled using a stochastic bi-level formulation, which compared with previous works added a new uncertain parameter: the uncertain rivals' offers to the day-ahead market. Dabbagh and Sheikh-El-Eslami [17] conducted risk-averse studies for a VPP (aggregating a wind power farm, a non-dispatchable load, a conventional power plant, a pumped hydro storage plant, and a flexible demand) using a risk-averse SP formulation.

In these works, the length of the time horizon is, naturally, 24 hours; given that the decision-making framework involves the participation in the day-ahead and balancing markets. Lima et al. [6] studied two-stage risk-averse SP approaches and suitable decomposition algorithms to solve a VPP (aggregating a thermal plant, a wind power farm, and a pumped-storage hydro plant) scheduling and electricity market

participation. They considered a time horizon of 168 hours and samples with sizes up to 25,000 elements. Consequently, in contrast with the works mentioned above that considered shorter horizons and smaller samples, they explored different decomposition algorithms to handle the CVaR in the objective function and large size SP formulations. Lima et al. [6] also proposed a wind power forecast methodology based on a wind speed ensemble obtained from a weather forecast model.

Recently, Castillo et al. [18] integrated a risk-averse SP formulation into a rolling horizon approach to optimize the operations of a VPP (aggregating a battery energy storage system, a photovoltaic system, a micro-grid, and a diesel generator). These authors concluded that for extreme weather conditions, the risk-averse solutions increase the value to the VPP.

1.2. Robust optimization formulations for VPPs

Lima et al. [19] proposed an adaptive robust optimization (ARO) formulation for the optimal scheduling of a VPP (aggregating a thermal plant, a wind power farm, and a pumped-storage hydro plant) considering a two-stage decision framework. Their model involves uncertainty in electricity prices and wind power production, and the formulation requires a decomposition algorithm to handle the resulting two-stage formulation. In that work, a detailed study on the computational performance of two decomposition algorithms and the analysis of the impact of the risk parameters on the results were performed. Shabanzadeh et al. [20] presented also a robust optimization formulation for a VPP, which considered the uncertainty on the electricity prices in the objective function. Their robust optimization formulation does not consider the two-stage framework as in Lima et al. [19], and therefore, it is solved directly as an MILP problem, without requiring a decomposition algorithm. Rahimiyan and Baringo [21] addressed the strategic bidding for a price-taker VPP (price-responsive demands, wind power plant, and an energy storage facility) in the day-ahead and real-time markets. Their formulation considers uncertainty in the wind power production and electricity prices. Compared to Lima et al. [19], they consider two types of markets, but do not consider the decisions on the forward contracts. The time horizon considered is 24 hours and an out-of-sample analysis is performed for seven consecutive days.

There is a recent line of research that encompasses in a single formulation the SP and ARO features, whereby, some uncertain parameters are described by discrete distributions and other parameters by uncertainty sets. For example, Baringo and Baringo [22] followed this approach for the offering strategy of a VPP (aggregating a conventional power plant, a wind power unit, a storage unit, and a flexible demand), where in their formulation the uncertainty in the wind power production is represented with uncertainty sets as in a robust optimization approach, and the uncertainty in electricity prices using a sample with discrete elements. As in Rahimiyan and Baringo [21], Baringo and Baringo [22] consider a time horizon of 24 hours and an out-of-sample analysis for four consecutive days. In other studies related with power systems, Zhang and Conejo [23] proposed a transmission expansion planning model that considers short-term and long-term decisions, where the uncertainty description is selected as a function of its nature. The long-term uncertainty parameters are described by robust uncertainty sets, whereas the short-term uncertainty parameters are represented by discrete scenarios. Fanzeres et al. [24], in the context of an electricity market, proposed a hybrid bi-level formulation with uncertainty sets for the electricity prices and a discrete distribution for the renewable generation. Another example in the context of power systems is the formulation proposed by Zhao and Guan [25], which merges an expectation and a worst-case measure in the objective function.

1.3. Comparison between risk-averse methodologies: SP vs ARO

One of the first comparisons between risk-averse SP and RO was made by Lagos et al. [26], involving VaR and CVaR within SP, and RO models for an open pit-mining problem without recourse. In the context of power systems, SP and ARO have been compared using unit commitment (UC) problems. For example, Pandzic et al. [27] compared four UC formulations under wind uncertainty with a focus on reliability. Two of these formulations are risk-neutral SP and ARO. They found that the ARO formulation leads to slightly higher optimal objective function values than the SP formulation, but the ARO formulation is generally computationally more efficient. [28] also contrasted four UC formulations under demand uncertainty: probabilistically constrained optimization, RO, two-stage SP, and two-stage ARO on real-world instances. As

the previous reference, a risk-neutral SP formulation is compared to an ARO formulation. Their results show that the SP formulation leads to comparatively less robust results with higher computational cost. Kazemzadeh et al. [29] focused on two UC formulations involving risk: a risk-averse SP formulation and an ARO formulation. They built two types of uncertainty sets for ARO, one based on ranges and another one that includes probabilities of scenarios together with ranges. They further analyze the trade-offs of cost vs. risk in solutions of the SP formulation and ARO formulation and conclude that for high conservatism, ARO leads to the best trade-off.

The three references above focus on a single problem, the UC, and two of them limit the comparisons to risk-neutral SP vs. ARO formulations. These studies raise relevant questions: 1) Do the computational trends remain valid for other problems? Are these trends general? 2) What is the difference in terms of actual solutions between risk-averse SP and ARO? 3) How does the type of data, in terms of parameters and uncertainty, affect the computational behavior and actual solutions?

We expand the conclusions of those works and answer the questions above by analyzing risk-averse SP and ARO formulations applied to a VPP problem, and using a systematic approach. We adopt a sample average approximation (SAA) methodology, and perform a comprehensive comparison that covers: a) computational performance depending on sample sizes and risk parameters; b) assessing approximate first-stage optimal solutions, obtained using the SAA methodology and using inference statistics on the expected profit and CVaR; and c) analyzing impact of the uncertainty characterization on the computational efficiency and results.

To make a fair comparison in terms of computational performance, the SP approach relies on an efficient implementation of the L-Shaped method that handles a risk measure in the objective function, based on [6] and the ARO approach relies on a tailored column and constraint generation (CCG) method, based on [19]. Compared to Lima et al. [19], we implement a CCG method embodying new convergence acceleration concepts. Therefore, adapting and extending previous contributions, the main contributions of this work are twofold: 1) providing a comprehensive and systematic comparison between two-stage risk-averse SP and two-stage ARO; and 2) deriving insights resulting from the analysis that, in particular for the case study considered, indicate that SP and ARO can lead to similar solutions, depending on the risk parameterization in each approach.

This paper is organized as follows: in Section 2, the VPP model and underlying assumptions are presented. In Section 3, a generic formulation of the VPP optimization problem under uncertainty is presented, and is exploited to derive the SP formulation, the ARO formulation, and the associated uncertainty representations. In Section 4, the two decomposition algorithms and the corresponding problem formulations are described. The computational results and risk analyses are presented and discussed in Section 5. We present the conclusions of the work in Section 6.

2. Problem statement

We address the self-scheduling and market involvement of a VPP using two alternative approaches: risk-averse SP and ARO. The VPP problem is defined in a general two-stage decision framework that fits the two approaches. The planning horizon considered is one week, divided in 168 time periods of one hour. Figure 1 illustrates the context of this work, the decision framework, and the constitution of the VPP. The VPP has been described in detail in Lima et al. [19] and Lima et al. [6]. The optimal management of the VPP involves a self-scheduling problem and the interaction with the electricity market. The self-scheduling problem determines the pump-storage hydro plant and thermal plant operation subject to their operational constraints. This means determining for each period a) the power output/input from/to the pump-storage; and b) the power output from the thermal plant. The electricity market interaction involves selecting forward contracts to sell or buy electricity through one or two contracts and the corresponding fixed power. There are two contracts offered to the VPP with a block structure, each block having a fixed price and power. In addition, the interaction involves the prescription of the power to sell or buy from the electricity market pool in each period.

We assume a two-stage decision framework. The first-stage decisions are made before the planning horizon: the selection of the characteristics of the contract and the commitment of the thermal unit; these

are the here-and-now decisions [30]. The second-stage decisions are made during the planning horizon: the dispatch of the thermal and pump-storage hydro plants, and the purchasing and selling of power in the electricity market in each period; these are the wait-and-see decisions. The first-stage decisions are fixed during the planning horizon, whereas the second-stage decisions depend on the realization of the uncertainty and are fixed in each period. The option to buy electricity from the pool implies that there are always feasible second-stage decisions for any first-stage decisions (in terms of the contract setup and the thermal commitment constraints).

The wind power profile and the electricity price profile are uncertain. The uncertainty in the wind power is characterized by an ensemble of 51 members derived from a wind speed ensemble from a weather forecast model; the forecast of the uncertain electricity prices uses a regression model. Based on these descriptions, the uncertainty treatment is further adapted according to the selected optimization approach.

The objective of the VPP is to maximize the operating profit that results from selling electricity, by forward contracts and in the electricity pool, minus the costs associated with the generation and purchases of electricity through the contracts or in the pool. The profit is uncertain given that it depends on the uncertain wind power and electricity prices. The full formulation of the deterministic problem and the data of operation of the generation units and contracts are presented in Appendix Appendix A (online supplemental material), and the uncertainty models of the wind power and electricity prices are outlined in Appendix Appendix B (online supplemental material).

In the problem considered, three important components interact: the time horizon of 168 hours, the risk-aversion level, and the use of forward contracts. On the one hand, the time horizon of 168 hours has the advantage of reducing shortsighted effects at the end of a time horizon of 24 hours when dealing with water storage and allows the pump-storage hydro unit to look forward to the full week. Therefore, the 168 hours enable better planning for using electricity and managing water. In addition, it allows a more effective planning of minimum up and down time constraints, ramp rates limits, and startup costs of the thermal

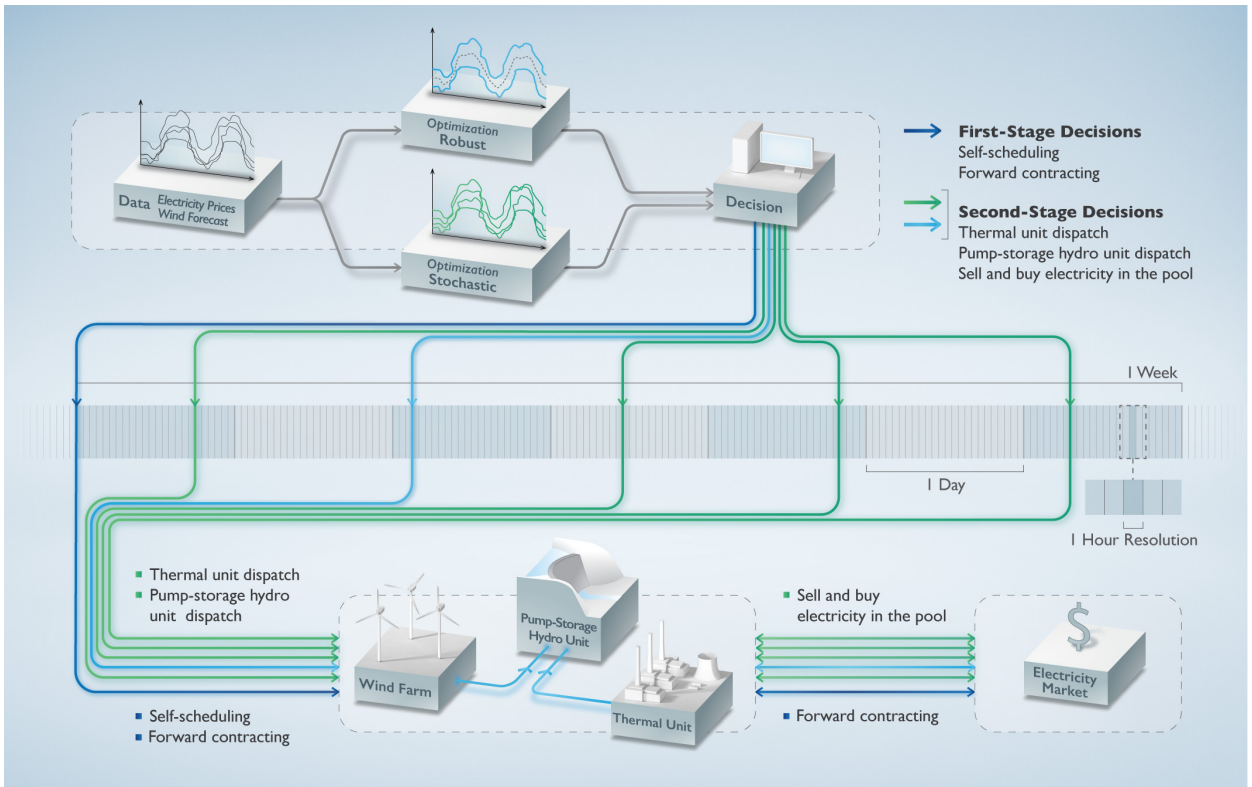


Figure 1: VPP composition and decision framework.

unit across the days of the week. On the other hand, the decision-makers need to comply with the contracts established, but the thermal unit commitment decisions may only be implemented for the first day of the week. Afterward, a similar model can be run during the week (on a rolling window) with updated forecasts of the electricity prices and wind power outputs to determine daily commitments. This procedure does not prevent the commitment decisions to be modeled as first-stage decisions to provide feedback information for the contract decisions. One way to improve the model would be considering the commitment variables as second-stage variables, or using a multi-stage decision framework.

The 168-hour time horizon allows considering weekly forward contracts. In general, these contracts can be established via an organized futures market or via bilateral contracts [4]. A number of works in the forward contracts literature identify weekly contracts as a tool to hedge against the volatility of electricity markets; see for example [31], Deng and Oren [32], Kristiansen [33], Benth and Koekebakker [34], Benth et al. [35], Botterud et al. [36], Kristiansen [33], and Weron and Zator [37], Aïd [38]. Specifically, Hope et al. [31] discussed the advantages of weekly contracts over yearly contracts, and Aïd [38] described electricity markets by country/region, including a reference to weekly contracts as an option in some markets. Considering daily bidding without contract selection misses an important tool for risk management, whereas the integration of weekly planning and contracts provides such a tool.

The proposed risk-aversion models provide instruments for decision-makers to mitigate low weekly profits (or losses). These instruments are valuable in situations that require risk mitigation. Decision-makers have different risk acceptance levels and may adopt different risk-averse approaches depending on their risk perception. For example, after one week of low-profit, a risk-neutral decision maker may reduce the risk acceptance level for the following week to mitigate a potential sequence of low profits. Therefore, in some circumstances, low weekly profits might be undesirable, and we investigate two approaches to manage that risk. The relevance of this argument increases if the planning horizon is increased from one week to one month. In addition, comparing alternative risk-averse solutions with risk-neutral ones provides estimates of expected profit reductions that can further support the decision-maker with risk acceptance levels. Additionally, an extreme risk-averse approach can be used in situations requiring only the minimization of risk and quantifying the advantages or disadvantages of a worst-case solution, which we discuss in detail in this work.

3. Risk-averse optimization models

In this section, we start by introducing a general and compact formulation for a stochastic optimization formulation that captures the main structural properties of the VPP problem: a) the two-stage decision process of the VPP; b) binary variables in the first-stage; c) uncertainty in the right-hand-side of constraints; and d) uncertainty in the objective function. This formulation aims to introduce the methods applied to the VPP problem rather than represent all the variables and constraints of the VPP problem. A detailed deterministic model for the VPP problem is presented in Appendix Appendix A (online supplemental material) and the corresponding stochastic model in Appendix Appendix C (online supplemental material). The formulation is as follows

$$
\begin{aligned}
w^* := \max_{x,z,y(\theta)} \quad & \psi(x, z, y(\theta)) \\
s.t. \quad & Ax + Bz \leq b \\
& Cz + Dy(\theta) \leq d, \quad \forall \theta \in \Theta \\
& Fy(\theta) + Gx = h(\theta), \quad \forall \theta \in \Theta \\
& x \in \mathbb{R}^{n_1}_+, z \in \mathbb{B}^{n_2}, y(\theta) \in \mathbb{R}^{n_3},
\end{aligned}
\tag{1}
$$

where $\mathbb{B} := \{0, 1\}$, x, z are deterministic vectors with dimensions n_1 and n_2, repsectively, $y(\theta)$ is a vector with size n_3, $A \in \mathbb{R}^{m_1 \times n_1}$, $B \in \mathbb{R}^{m_1 \times n_2}$, $C \in \mathbb{R}^{m_2 \times n_2}$, $D \in \mathbb{R}^{m_2 \times n_3}$, $F \in \mathbb{R}^{m_3 \times n_3}$, $G \in \mathbb{R}^{m_3 \times n_1}$ are deterministic matrices with known coefficients, and $b \in \mathbb{R}^{m_1}$, $d \in \mathbb{R}^{m_2}$ are deterministic vectors with known parameters. The random event θ belongs to the set Θ of future events. The vector $y(\theta)$ is optimally determined for each

event and the constraints are satisfied for any event. The objective function in (1) to be maximized is of the form

$$\psi(x, z, y(\theta)) := \rho\left[f(x, z, y(\theta), \theta)\right], \tag{2}$$

where ρ is a functional (such as the CVaR or expectation) and f is defined as $f(x, z, y(\theta), \theta) := c^\intercal x + \bar{c}^\intercal z + \tilde{c}^\intercal(\theta) y(\theta)$, with $c \in \mathbb{R}^{n_1}$ and $\bar{c} \in \mathbb{R}^{n_2}$ denoting deterministic vectors with known parameters.

To bridge the formulation in (1) with the VPP problem, we establish the connection between the variables of both in the remaining of this paragraph. The function f represents the difference between the revenues of selling electricity minus the costs of generation and buying electricity. The vector $c^\intercal$ represents the prices related with contracts, $\bar{c}^\intercal$ denotes the fixed costs related with the commitment of the considered thermal unit, $\tilde{c}^\intercal(\theta)$ captures the uncertain prices of the electricity in the pool market. The vectors x and z are the first-stage variables, where x represents the power to sell or buy through contracts and z the binary variables associated with the commitment of the thermal unit and the selection of the blocks of the contracts. The vector y represents the second-stage variables, which stands for the output generation of the thermal unit, the output generation and consumption of the pump-storage hydro plant and the corresponding volumes of water in the reservoir and flows of water, and the energy to sell or buy in the electricity pool market. The random vector $h(\theta)$ represents the wind power.

The first set of constraints in Problem (1) captures the structure of the contracts and the constraints on the commitment of the thermal unit. The second set of constraints represents the limits of operation of the thermal unit (maximum and minimum power output and power output ramp rates limits), and the limits of operation of the hydro plant (maximum and minimum power output and limits on the flows of water and volume of water in the reservoir). The last constraint expresses the energy balance of the VPP.

The definition of the profit functional ρ and the treatment of the uncertainty differentiate SP and ARO formulations, whose derivations are detailed in the next two subsections.

3.1. Risk-averse stochastic programming formulation

Problem (1) can be recast as a SP problem considering that θ belongs to a probability space $(\Theta, \mathcal{F}, P)$, with Θ being the set of future events, $\mathcal{F}$ a σ-algebra, and P a probability measure. We consider a risk-averse SP formulation that incorporates an objective function involving the expectation of the profit and the CVaR of the profit. The CVaR of the profit f pertaining to the $(1-\alpha)$-quantile is the conditional expectation of f given that f does not exceed the value at risk (VAR) of f. The VaR of f for the $(1-\alpha)$-quantile is defined as $\text{VaR}_{1-\alpha}[f] = \max\{v | F_f(v) \leq 1-\alpha\}$, where F_f is the cumulative distribution function of f; and the CVaR is defined as $\text{CVaR}_{1-\alpha}[f] = \mathbb{E}\left[f | f \leq \text{VaR}_{1-\alpha}[f]\right]$.

Denoting $\eta = \text{VaR}_{1-\alpha}[f]$ and applying the CVaR to the profit function f, it can be shown that $\max_{x,z,y(\theta)} \left\{\text{CVaR}_{1-\alpha}[f(x, z, y(\theta), \theta)]\right\} = \max_{x,z,y(\theta),\eta} \left\{\mathbb{E}\left[\eta - \frac{1}{1-\alpha}\big(\eta - f(x, z, y(\theta), \theta)\big)^+\right]\right\}$. These definitions are given for completeness and we refer to Pflug [39], Rockafellar and Uryasev [40], and Rockafellar [41] for the properties of the CVaR and its reformulation used in stochastic programming.

Therefore, the SP formulation is driven by the maximization of a combination of the expectation and the CVaR of the profit:

$$\max_{x,z,y(\theta)} \psi[f(x, z, y(\theta), \theta)] := \max_{x,z,y(\theta),\eta} \left\{\mathbb{E}\left[(1-\beta) f(x, z, y(\theta), \theta) + \beta\left(\eta - \frac{1}{1-\alpha}(\eta - f(x, z, y(\theta), \theta))^+\right)\right]\right\}, \tag{3}$$

where $\beta \in [0, 1]$ defines the weights of the expectation and CVaR of f, $\alpha \in [0, 1]$ defines a quantile, and $(\cdot)^+$ represents $\max\{\cdot, 0\}$. The resulting formulation of (1) with (3) can be recast into a canonical two-stage formulation to fit the VPP two-stage decision process. The objective function in (3) is parameterized over both β and α. These parameters translate the levels of risk that the decision-maker is willing to accept, and thus, drive the formulation to more or less conservative solutions. The formulations defined with (3) were studied in Lima et al. [6], where the focus was on the efficient solution of these formulations with an adapted L-Shaped algorithm to handle the CVaR.

In comparing with the ARO approach, we focus on maximizing the CVaR of the profit in (3) with $\beta = 1$. In this case, the CVaR of the first-stage profit is given by $c^\intercal x + \bar{c}^\intercal z$ due to the translation invariance property

of CVaR [42, 39]. The objective function is re-written as $\psi\left(x,z,y(\theta)\right) := c^{\intercal}x + \overline{c}^{\intercal}z + \mathcal{Q}(x,z)$, where the CVaR of the second-stage profit is given by $\mathcal{Q}(x,z) := \text{CVaR}_{1-\alpha}[Q(x,z,\theta)] = \mathbb{E}\Big[\eta - \dfrac{1}{1-\alpha}[\eta - Q(x,z,\theta)]^{+}\Big]$, which leads to the two-stage canonical SP formulation:

$$\begin{aligned} w^{*} := \max_{x,z} \quad & c^{\intercal}x + \overline{c}^{\intercal}z + \text{CVaR}_{1-\alpha}[Q(x,z,\theta)] \\ s.t. \quad & Ax + Bz \leq b \\ & x \in \mathbb{R}^{n_1}_{+}, z \in \mathbb{B}^{n_2}, \end{aligned} \tag{4}$$

with

$$\begin{aligned} Q(x,z,\theta) := \max_{y(\theta)} \quad & \tilde{c}^{\intercal}(\theta)y(\theta) \\ s.t. \quad & Dy(\theta) \leq d - Cz, \\ & Fy(\theta) = h(\theta) - Gx, \\ & y(\theta) \in \mathbb{R}^{n_3}. \end{aligned} \tag{5}$$

3.2. Robust optimization formulation

The conversion of Problem (1) into an ARO counterpart problem [12] is based on two ideas: 1) a particular treatment of the support of the random vectors h and $\tilde{c}$ using uncertainty sets; and 2) finding a robust solution for the first-stage variables x and z. It is called robust because it aims at maximizing f in the "worst" random event in the support (range) of the random events. The objective function of the ARO counterpart problem of (1) is $\psi_{RO}\left(x,z,y(\theta)\right) := \inf_{\theta\in\Theta} f(x,z,y(\theta),\theta)$. It leads to the ARO counterpart formulation:

$$\begin{aligned} w^{*}_{RO} := \max_{x,z,y(\theta)} \quad & \inf_{\theta\in\Theta}\left\{c^{\intercal}x + \overline{c}^{\intercal}z + \tilde{c}^{\intercal}(\theta)y(\theta)\right\} \\ s.t. \quad & Ax + Bz \leq b \\ & Cz + Dy(\theta) \leq d, \quad \theta \in \Theta \\ & Fy(\theta) + Gx = h(\theta), \quad \theta \in \Theta \\ & x \in \mathbb{R}^{n_1}_{+}, z \in \mathbb{B}^{n_2}, y(\theta) \in \mathbb{R}^{n_3}. \end{aligned} \tag{6}$$

Problem (6) can be re-formulated into a two-stage decision process with first-stage variables x and z, and second-stage variables y, by defining $\psi_{RO}\left(x,z\right) := c^{\intercal}x + \overline{c}^{\intercal}z + \inf_{\theta\in\Theta} Q(x,z,\theta)$ which leads to the following two-stage ARO formulation:

$$\begin{aligned} w^{*}_{RO} := \max_{x,z} \quad & c^{\intercal}x + \overline{c}^{\intercal}z + \inf_{\theta\in\Theta} Q(x,z,\theta) \\ s.t. \quad & Ax + Bz \leq b \\ & x \in \mathbb{R}^{n_1}_{+}, z \in \mathbb{B}^{n_2}, \end{aligned} \tag{7}$$

with $Q(x,z,\theta)$ defined as in (5), but with a distinct support of θ. Formulations (4) and (7) represent the two-stage decision process, where in the first stage the forward contracts and the thermal unit commitment are derived to maximize the first-stage profit $c^{\intercal}x + \overline{c}^{\intercal}z$. The second-stage decisions—dispatch of the thermal unit and pump-storage hydro plant and interaction with the electricity market—are optimally selected to maximize the CVaR of the second-stage profit $\tilde{c}^{\intercal}(\theta)y(\theta)$ in (4) and the second-stage profit $\tilde{c}^{\intercal}(\theta)y(\theta)$ for the worst random event of the uncertain parameters in (7).

The definition and construction of uncertainty sets from data is a key research area in ARO. Advanced approaches to define the uncertainty sets involve hypothesis testing or the utilization of risk measures [43, 44]. The uncertainty sets usually include at least two components: 1) the range of the random vectors; and 2) a risk management approach. The first defines only the range of the random vectors without considering any probability distribution over this range, whereas the risk management approach is used to avoid overly conservative solutions due to the worst-case approach.

In this work, the support of $h(\theta)$ and $\tilde{c}(\theta)$, which in our problem represent the uncertainty of the wind power and electricity prices, are specified using a discretization of convex uncertainty sets. For instance, the uncertainty set for h is written as

$$\mathcal{U}_h := \left\{ h : h = h^f + \operatorname{diag}\left(u^+\right) h^u - \operatorname{diag}\left(u^-\right) h^l, u^+, u^- \in \mathbb{B}^{m_3} \right\}, \tag{8}$$

where h^f is the forecast value, h^u, h^l are upper and lower deviations, respectively, u^+, u^- are binary random vectors, and $\operatorname{diag}(u)$ is a diagonal matrix defined by the vector u. Therefore, the uncertainty sets are built around forecast values h^f and $\tilde{c}^f$ using bounded intervals. We refer the reader to Ben-Tal et al. [12] for a detailed description on the construction of this type of uncertainty set using perturbation vectors and parameterizations, as well as analysis on the tractability of robust counterpart problems that use convex uncertainty sets.

The main difference here with the sets proposed in Ben-Tal et al. [12] is the introduction of the binary variables in (8), which has the advantage of keeping the subproblem within the scope of the decomposition algorithm used to solve the robust counterpart as an MILP problem; see Thiele et al. [45], Jiang et al. [46], and Lima et al. [19].

The conservatism of an ARO formulation can be further controlled introducing a budget of uncertainty constraint in the definition of the uncertainty sets [47]. The budget of uncertainty constraint limits the deviations of the elements of the random vectors h and $\tilde{c}$ from their forecast values h^f and $\tilde{c}^f$. The uncertainty set in (8) is augmented by introducing a budget of uncertainty constraint through $e^\intercal \left(u^+ + u^-\right) \leq \Gamma$, where e is a vector with all entries equal to one, and $\Gamma > 0$ is a known parameter defined by the decision-maker, which represents the budget of uncertainty that the decision-maker is willing to account for. In this work, Γ controls the number of time periods where the uncertain prediction is allowed to depart from the corresponding forecast values. Note that more general uncertainty budget constraints may be defined with varying importance of the entries of u as suggested in Bertsimas and Sim [48]. Tuning Γ appropriately allows to control the deviations from the forecast value at the end of the time horizon. However, larger deviations at the end of the time horizon can also be captured in the range of the random vectors. This is the case in this work, where the variance of the wind power and electricity prices increases in time. With this setup, the ARO formulation presented later with $\Gamma = 0$ is equivalent to the expected value problem [30], where the random vectors are replaced by their mean value; with $\Gamma = \Gamma_{\max}$, where $\Gamma_{\max}$ is the maximum number of deviations, the ARO formulation recasts the conservative robust approach, where the worst random event is considered.

In the literature, the uncertainty sets used in applications are usually defined in one of the following ways: 1) one interval range for each random component of the vectors and one budget of uncertainty constraint per random component [8, 46, 19, 49, 50]; 2) one interval range for each random component and two budget of uncertainty constraints per random component [51, 25]; 3) one interval range for each random component, one budget constraint, and a correlation between random components [52]; or 4) one interval range for each random component, one budget constraint, and additional constraints that limit the variations of the random component between consecutive periods of time [24]. The limits of the interval range for the random variable are usually defined as a percentage of the forecast value, or as a given percentile of the forecast prediction [8, 25]. In addition, Lorca and Sun [53] proposed dynamic uncertainty sets to describe the wind power uncertainty from wind farms.

The final formulations of the uncertainty sets for h and $\tilde{c}$ are the following

$$\mathcal{U}_h := \left\{ h : h = h^f + \operatorname{diag}\left(u^+\right) h^u - \operatorname{diag}\left(u^-\right) h^l, e^\intercal \left(u^+ + u^-\right) \leq \Gamma, u^+, u^- \in \mathbb{B}^{m_3} \right\}, \tag{9}$$

and

$$\mathcal{U}_{\tilde{c}} := \left\{ \tilde{c} : \tilde{c} = \tilde{c}^f + \operatorname{diag}\left(u^+\right) \tilde{c}^u - \operatorname{diag}\left(w^-\right) \tilde{c}^l, e^\intercal \left(w^+ + w^-\right) \leq \Gamma, w^+, w^- \in \mathbb{B}^{n_3} \right\}. \tag{10}$$

In Section 4, we provide more details on the ARO problem as well as on the solution method used.

3.3. The relation between the worst-case in SP and ARO

In the CVaR definition, for any well behaved f, we have two limiting cases [41]: 1) $\lim_{\alpha \to 0} \operatorname{CVaR}_{1-\alpha}[f] = \mathbb{E}[f]$; and 2) $\lim_{\alpha \to 1} \operatorname{CVaR}_{1-\alpha}[f] = \inf_{\theta \in \Theta}[f]$. The first equation shows that for $\beta = 1$ and $\alpha \to 0$, we obtain

the same solution as with $\beta = 0$. The second equation reveals a natural link between the SP objective function with $\beta = 1$, $\alpha \to 1$ and $\psi_{RO}(x, z, y(\theta))$. Both objective functions aim for the maximization of the worst-case profit. This interesting link between risk-averse SP and ARO was identified by [54].

As customary when dealing with finite numbers of samples, we shall refer by $\alpha = 1$ the worst-case approach, although in the computational implementation $\alpha = 1$ corresponds to $\alpha = (N-1)/N$ (which is close but not exactly 1 for large N).

3.4. Comparing solutions from the SP and ARO

SP and ARO rely on fundamentally different representations of uncertainty, which favor one approach vs. the other depending on the problem considered. If the probability distribution of the random variables is not known, an ARO approach is the obvious choice. On the other hand, if a probability distribution is available, then SP is the natural choice. We address a case study where both approaches can be used, and provide insights on the overall comparison of risk-averse SP and ARO. Selecting the CVaR as the risk measure in SP is a natural choice due to its properties [40, 39, 41] and its direct link with ARO [54, 41].

The two approaches incorporate different objective functions, driven by different risk parameters: SP uses β and α, while ARO uses Γ. Therefore, the comparison of the optimal values of the objective function of each formulation needs to take these parameters into account. The optimal objective function values of the extreme risk-averse SP and ARO formulation correspond to the worst profit condition, and they do not provide additional information regarding profit for other uncertainty realizations. Note that in practice, the profit realization depends on the realization of the uncertain parameters and the optimal first-stage variables.

To compare solutions obtained from different formulations and risk parameterizations, we compare first-stage variables and estimates and confidence intervals of the expected profit and CVaR of the profit. These estimates and confidence intervals are calculated using the bound estimation stage of the SAA methodology, which it is applied to both the SP and ARO.

3.4.1. The SAA methodology.

To compute solutions of the SP problems and assess their quality, we use the SAA methodology developed in Lima et al. [55]. For the sake of completeness, we briefly outline its main concepts below. In the SSA methodology, estimates and confidence intervals for upper and lower bounds on the true optimal value are calculated, as well as an estimate of an upper bound on the gap between the true optimal value and the optimal value of each first-stage solution found.

The SP formulation with the objective functions defined in (3) fits the general formulation

$$w^* = \max_{x,z,y(\theta)\in\mathcal{W}} \{\mathbb{E}\left[\phi(x, z, y(\theta), \theta)\right]\}, \tag{11}$$

where $\mathcal{W} := \{x, z, y(\theta) | Ax + Bz \leq b; Cz + Dy(\theta) \leq d,\ \theta \in \Theta; Fy(\theta) + Gx = h(\theta),\ \theta \in \Theta; x \in \mathbb{R}^{n_1}_+, z \in \mathbb{B}^{n_1}, y(\theta) \in \mathbb{R}^{n_1}_+\}$, with the optimal value w^* and optimal solution x^*, y^*, z^*. [56] proposed a solution approach for Problem (11) using the approximation problem

$$w^*_{\mathcal{N}} = \max_{x,z,y(\theta)\in\mathcal{W}_{\mathcal{N}}} \Big\{\frac{1}{N}\sum_{n=1}^{N} \phi(x, z, y^n, \theta^n)\Big\}, \tag{12}$$

for a sample set of random events θ^n. We refer to Problem (12) as the sample average approximation problem. In (12), $\mathcal{N}$ denotes a sample with $N = |\mathcal{N}|$ elements. The solution of Problem (11) is then approximated by the optimal value and solution of Problem (12), namely $w^*_{\mathcal{N}}$ and $x^*_{\mathcal{N}}$, $y^*_{\mathcal{N}}$, and $z^*_{\mathcal{N}}$. SAA involves two stages: an optimization stage and a bound estimation stage.

The optimization stage involves solving Problem (12) for independent samples with the same size, whereas the bound estimation stage involves selecting and fixing a first-stage solution in Problem (12), that is $x^*_{\mathcal{N}}$, $z^*_{\mathcal{N}}$, and solving it for multiple samples with the same size. The bound estimation stage determines an estimate with confidence interval for a lower bound on the optimal value. However, in the context of this

work, we are interested in additional results associated with the lower bound on the true optimal solution. More specifically, we seek estimates and confidence intervals on the expected profit and CVaR of the profit of the VPP. We refer to Lima et al. [55] for further details on the implementation of the SAA methodology adopted here. A central part of SAA is the solution of the Problem (12) for a given sample. In Section 4, we describe the method used to solve this problem.

4. Solution methods

In this section, we provide additional details about the optimization formulations and the solution methods used in SP and ARO.

4.1. Stochastic programming decomposition algorithm

The SP formulation is solved with the L-Shaped algorithm [57]. In particular, to tackle the risk-averse formulations, we use a single-cut version of an adapted L-Shaped algorithm; see Lima et al. [6]. In that work, the superior performance of an adapted L-Shaped algorithm as compared with the solution of the extensive formulation was demonstrated. The master problem is defined as

$$
\begin{aligned}
w^*_{\beta\mathcal{N},\mathcal{M}} = \max_{x,z,\omega_1,\omega_2} \quad & c^\intercal x + \bar{c}^\intercal z + (1-\beta)\omega_1 + \beta\omega_2 \\
s.t. \quad & Ax + Bz \leq b \\
& \omega_1 \leq \sum_{n\in\mathcal{N}} p_n \left[(d - Cz)^\intercal \mu_n^m + (h_n - Gx)^\intercal \lambda_n^m\right], \quad \forall m \in \mathcal{M} \\
& \omega_2 \leq \sum_{n\in\mathcal{N}} p_n \Big(\eta^m - \frac{1}{1-\alpha} v_n^m\Big), \quad \forall m \in \mathcal{M},\, \beta > 0 \\
& v_n^m \geq \eta^m - \left[(d - Cz)^\intercal \mu_n^m + (h_n - Gx)^\intercal \lambda_n^m\right], \quad \forall n \in \mathcal{N}, \forall m \in \mathcal{M},\, \beta > 0 \\
& v_n^m \geq 0, \quad \forall n \in \mathcal{N}, \forall m \in \mathcal{M},\, \beta > 0 \\
& x \in \mathbb{R}^{n_1}, z \in \mathbb{B}^{n_2}, \omega_1 \in \mathbb{R},\, \omega_2 \in \mathbb{R},
\end{aligned}
\tag{13}
$$

and the dual subproblem as

$$
\left.
\begin{aligned}
\min_{\mu_n,\lambda_n} \quad & \left(d - Cz^k\right)^\intercal \mu_n + \left(h_n - Gx^k\right)^\intercal \lambda_n \\
s.t. \quad & D^\intercal \mu_n + F^\intercal \lambda_n \geq \tilde{c}_n \\
& \mu_n \in \mathbb{R}^{m_2}_+,\ \lambda_n \in \mathbb{R}^{m_3}
\end{aligned}
\right\} \quad \forall n \in \mathcal{N}, \tag{14}
$$

where μ_n and λ_n are the dual variables corresponding to n-th sample.

The adapted L-Shaped method relies on a master problem with one additional optimality cut, the third constraint in (13), and one extra variable, w_2, as compared to a traditional one. The master problem involves only optimality cuts. This is so because, for the formulation considered, each first-stage solution leads to a feasible second-stage solution. This means that the formulation has relatively complete recourse. The CVaR definition involves two additional constraints per iteration and scenario (the fourth and fifth constraints) and one new variable, v_n^m, per iteration and scenario as well. The new optimality cut outer-approximates the entire CVaR term in the objective function, setting the VaR, denoted by η^m, as a second-stage variable. The VaR is calculated at each iteration after the solution of the LP subproblems and is used in the CVaR optimality cut and in the calculation of the lower bound. These details are presented in Algorithm 1.

4.2. Robust optimization decomposition algorithm

The ARO formulation cannot be solved directly due to its bilevel structure, and it requires a decomposition algorithm. One approach is to use a Benders Decomposition based algorithm, such as described in

Lima et al. [19]. In this algorithm, the master problem is defined as

$$\begin{aligned} w^*_{RO,\mathcal{M}} := \max_{x,z,\xi} \quad & c^\intercal x + \overline{c}^\intercal z + \xi \\ s.t. \quad & Ax + Bz \leq b \\ & \xi \leq (d - Cz)^\intercal \mu^m + (h - Gx)^\intercal \lambda^m, \quad m \in \mathcal{M} \\ & x \in \mathbb{R}^{n_1}_+, z \in \mathbb{B}^{n_2}, \xi \in \mathbb{R}. \end{aligned} \tag{15}$$

A careful inspection of the master problems associated with the risk-neutral SP and ARO, problems (13) and (15), shows that the master problems resemble each other. There are similarities because they are both derived based on the Benders Decomposition. The main difference between the two master problems is that, in the SP master problem, the optimality cut is built over the dual variables of all subproblems, whereas the optimality cut is only built over the dual variables of one subproblem in ARO. Therefore, in the L-Shaped algorithm for the risk-neutral formulation only one constraint per iteration is added to the master problem, but this constraint has a number of terms proportional to the sample size. The master problems that consider the CVaR in the objective function require additional constraints and variables per iteration, compared to the risk-neutral master problem (see Table 1), namely to outer-approximate the CVaR term in the objective function. Note also that, due to the relatively complete recourse property of the VPP formulation, problem (15) only incorporates optimality cuts.

In this work, we use a primal version of the master problem [8, 58]:

$$\begin{aligned} w^*_{RO,\mathcal{M}} := \max_{x,z,\xi} \quad & c^\intercal x + \overline{c}^\intercal z + \xi \\ s.t. \quad & Ax + Bz \leq b \\ & \xi \leq \tilde{c}^{\intercal m} y^m, \quad m \in \mathcal{M} \\ & Cz \leq d - Dy^m, \quad m \in \mathcal{M} \\ & Gx = h^m - Fy^m, \quad m \in \mathcal{M} \\ & x \in \mathbb{R}^{n_1}_+, z \in \mathbb{B}^{n_2}, y^m \in \mathbb{R}^{n_3}_+, \xi \in \mathbb{R}. \end{aligned} \tag{16}$$

We rely on the primal variant because for the same $\tilde{c}^{Tm}$ and h^m, the optimal value of the objective function of (15) is an overestimation of that of (16) [58, Proposition 3]. Besides, the computational performance of the primal version is better than or equal to the dual variant; see Zeng and Zhao [58], Lima et al. [19], and a comparison for the case studies in this work in Appendix Appendix D (online supplemental material). Note that a similar strategy can be implemented in the SP by transferring to the master problem some constraints of the primal subproblems.

The dual subproblem in the ARO decomposition algorithm is given by

$$\begin{aligned} \min_{\mu,\lambda,w^+,w^-,u^+,u^-} \quad & \left(d - Cz^k\right)^\intercal \mu + \left(h^f + \operatorname{diag}\left(u^+\right) h^u - \operatorname{diag}\left(u^-\right) h^l - Gx^k\right)^\intercal \lambda \\ s.t. \quad & D^\intercal \mu + F^\intercal \lambda \geq \tilde{c}^f + \operatorname{diag}(w^+)\tilde{c}^u - \operatorname{diag}(w^-)\tilde{c}^l \\ & e^\intercal \left(u^+ + u^-\right) \leq \Gamma \\ & e^\intercal \left(w^+ + w^-\right) \leq \Gamma \\ & \mu \in \mathbb{R}^{m_2}_+, \ \lambda \in \mathbb{R}^{m_3}, w^+, w^- \in \mathbb{B}^{n_3}, u^+, u^- \in \mathbb{B}^{m_3}. \end{aligned} \tag{17}$$

Note that the problem (17) has two bilinear terms in the objective function: $\operatorname{diag}\left(u^+\right)\lambda$ and $\operatorname{diag}\left(u^-\right)\lambda$. These specific bilinear terms involving a binary and a continuous variable result from using binary variables in the definition of the uncertainty set. This type of bilinear terms can be linearized as described in Lima et al. [19].

4.3. An adaptive maximum solution time to solve subproblems

The ARO algorithm involves the solution of two MILP problems on each iteration: the master problem (16) and the subproblem (17), using a branch & cut method. Preliminary computational experiments

Algorithm 1 L-Shaped algorithm adapted to handle the master problem (13).
Sets:= $\mathcal{N}$ - elements of one sample, $\mathcal{K}$ - iterations, $\mathcal{M}$ - optimality cuts.

Initialization: $LB = -\inf$, $UB = +\inf$, $k := 0$, $m := 0$, $\mathcal{M} := \emptyset$
while $|UB - LB| / |LB| > \epsilon$ **do**
  Solve the master problem, for $k = 0$, set $\omega_1 = 0$

Problem (13)

  Let $(x^k, z^k, \omega_1^k, \omega_2^k)$ be the optimal solution of the master problem
  Calculate the upper bound

$$UB = \min \left\{ UB, c^\intercal x^k + \bar{c}^\intercal z^k + (1-\beta)\,\omega_1^k + \beta \omega_2^k \right\} \tag{18}$$

  Solve the N LP subproblems

Problem (14)

  Let (λ_n^k, μ_n^k) be the optimal solution of the subproblem n
  Add cuts to the master problem built from the dual variables
  $m := m + 1$, $\mathcal{M} := \mathcal{M} \cup m$

$$\lambda_n^m = \lambda_n^k, \quad \forall n \in \mathcal{N}, \qquad \mu_n^m = \mu_n^k, \quad \forall n \in \mathcal{N} \tag{19}$$

  Calculate the η^m for the α-quantile of the distribution of $Q\left(x^k, z^k, \xi_n\right)$, $\forall n \in \mathcal{N}$
  Add optimality cuts to the master problem
  Calculate the $\text{CVaR}_{1-\alpha}^m$ of the distribution of $Q\left(x^k, z^k, \xi_n\right)$, $\forall n \in \mathcal{N}$

$$\text{CVaR}_{1-\alpha}^m = \eta^m - \frac{1}{1-\alpha} \sum_{n \in \mathcal{N}} p_n \left(\eta^m - \left[\left(d - Cz^k \right)^\intercal \mu_n^m + \left(h_n - Gx^k \right)^\intercal \lambda_n^m \right] \right)^+ \tag{20}$$

  Calculate the lower bound

$$LB = \max \left\{ LB, c^\intercal x^k + \bar{c}^\intercal z^k + (1-\beta) \sum_{n \in \mathcal{N}} \left\{ p_n \left[\left(d - Cz^k \right)^\intercal \mu_n^m + \left(h_n - Gx^k \right)^\intercal \lambda_n^m \right] \right\} + \beta \text{CVaR}_{1-\alpha}^m \right\} \tag{21}$$

end while

suggested that a) the master problem is relatively faster to solve than the subproblems; and b) in some cases, the subproblem solution took practically the total computational time assigned to the algorithm, which made the algorithm to terminate after a few iterations with a poor solution. These results indicate that limiting the solution time of the subproblems during the early iterations might be advantageous.

To enable the algorithm to overcome the low performance of the subproblems in early iterations, we use an adaptive maximum solution time strategy to solve the subproblems. The strategy is based on initially allocating a small maximum computational time for the solution of the subproblems. Then if the ARO algorithm does not improve the upper bound between two iterations, then the maximum time for the MILP subproblem is increased. On the other hand, the termination of the subproblem solution due to a time limit implies that the subproblem solution may be suboptimal. Therefore, to ensure the validity of the lower bound of the algorithm, this bound is calculated using the final lower bound on the objective function of the subproblem, obtained from the branch & cut method. Note that the computational performance of the ARO algorithm using this strategy depends on the effort put on the solution of the subproblem. The ARO algorithm's performance is discussed in Appendix Appendix E (online supplemental material) and additional performance results for alternative maximum computational times are discussed in Appendix Appendix F (online supplemental material). Here, we implement Algorithm 2.

4.4. Sizes of the master problems and subproblems

In Table 1, we present the size of the master problems associated with the SP and ARO decomposition methods, based on the compact and general formulation in (1). In SP, the size of the subproblems does not change with the number of iterations, but the number of subproblems to solve depends on the sample size N. In our SP setup, the subproblems are LP problems with n_3 constraints and $m_2 + m_3$ continuous variables. In ARO, there is only one subproblem to solve, but it is an MILP problem with $n_3 + 2$ constraints, $m_2 + m_3$ continuous variables, and $2n_3 + 2m_3$ binary variables. In our implementation of the L-Shaped method, the algorithm deals with the solution of N LP problems in parallel, whereas in ARO, the algorithm deals with one MILP, which in fact, involves the solution of a collection of LP problems within a branch & cut method. Ultimately, both approaches solve a collection of LP problems.

Algorithm 2 Primal constraint generation algorithm using the master problem (16).
Sets:= $\mathcal{K}$ - iterations, $\mathcal{M}$ - primal cuts

Initialization: $LB = -\infty$, $UB = +\infty$, $k := 0$, $\mathcal{M} := \emptyset$, $T^{\max \text{CPU}} = T_1$
while $|UB - LB| / LB > \epsilon$ **do**
  Solve the master problem, for $k = 0$ set $\xi = 0$

$$\text{Problem (16)}$$

  Let (x^k, z^k, ξ^k) be the optimal solution of the master problem
  Calculate the upper bound

$$UB = \min \left\{ UB, c^\intercal x^k + \overline{c}^\intercal z^k + \xi^k \right\}$$

  if $UB = c^\intercal x^{k-1} + \overline{c}^\intercal z^{k-1} + \xi^{k-1}$ **then**
    $T^{\max \text{CPU}} = T_2$
  end if
  Solve MILP subproblem with a maximum time limit of $T^{\max \text{CPU}}$

$$\text{Problem (17)}$$

  Let $\left(\lambda^k, \mu^k, w^{+k}, w^{-k}, u^{+k}, u^{-k}\right)$ be the optimal solution of the subproblem
  Let LB^{sp} be the lower bound on the MILP subproblem objective function obtained from the branch & cut method
  Set $m := m + 1$, $\mathcal{M} := \mathcal{M} \cup m$ and calculate h^m and $\tilde{c}^m$ based on (9) and (10)
  Add the primal cuts to the master problem
  Calculate the lower bound

$$LB = \max \{LB, LB^{sp}\}$$

end while

Table 1: Number of variables and constraints added to the master problems per iteration. The master problem has n_1 continuous variables, n_1 binary variables, and m_1 constraints, plus the numbers presented.

Formulation	Master problem	Variables	Constraints
SP - $\beta \in]0.0, 1.0[$	(13)	$2 + N/Iteration$	$(2 + N)/Iteration$
SP - $\beta = 0.0$	(13)	1	$1/Iteration$
SP - $\beta = 1.0$	(13)	$1 + N/Iteration$	$(1 + N)/Iteration$
ARO - Primal	(16)	$1 + n_3/Iteration$	$(1 + m_2 + m_3)/Iteration$

5. Computational experiments

In this section, we address the VPP problem by applying the risk-averse SP and the ARO. Our primary objective is to compare both approaches from the computational point of view and from the risk management perspective. Based on previous studies [19, 6], we use two efficient implementations of the decomposition algorithms for SP and ARO, and we assess their performance on two case studies taking into consideration different sample sizes and different risk management parameters. The SP decomposition algorithm, as well as the SAA methodology, rely on parallelization of the LP subproblems solution to reduce wall clock time. From the risk management perspective, we emphasize the performance analysis of the two approaches for the extreme risk-averse case, which establishes the link between risk-averse SP and ARO.

We consider a VPP consisting of a thermal unit, a wind farm, and a pump-storage hydro plant. Based on these units, we define two case studies: Case 1 and Case 2. The two cases assume the same wind power forecasts, same electricity price forecasts, and that the same contracts are offered to the VPP.

The main difference between Case 1 and Case 2 is the thermal unit, namely the performance characteristics that define the maximum power output, minimum up and down times, limits on the ramp rates, and the electricity generation costs. The thermal unit in Case 2 has lower maximum power output, minimum up and down times, and the electricity generation costs are higher than the ones of the thermal unit in Case 1. More specifically, the data for the thermal units are the following: the capacities (maximum power outputs) are 455 MW and 55 MW, the fixed costs are 1000 €/h and 660 €/h, and the variable generation costs are 16.19 €/MWh and 25.92 €/MWh for Case 1 and 2, respectively. In Case 2, the variable generation costs are higher, and the capacity is considerably lower than in Case 1, which leads to a ratio between the fixed cost and the capacity of 12 €/(hMW) for the thermal unit in Case 2 vs. 2.20 €/(hMW) in Case 1. In the discussion of the results, the term higher generation costs of the thermal unit in Case 2 takes also in consideration the impact of the lower capacity, by comparison with Case 1. The data and performance equations that characterize the region of operation of the generation units are described in detail in Appendix Appendix A (online supplemental material).

5.1. Uncertainty characterization for wind power and electricity prices

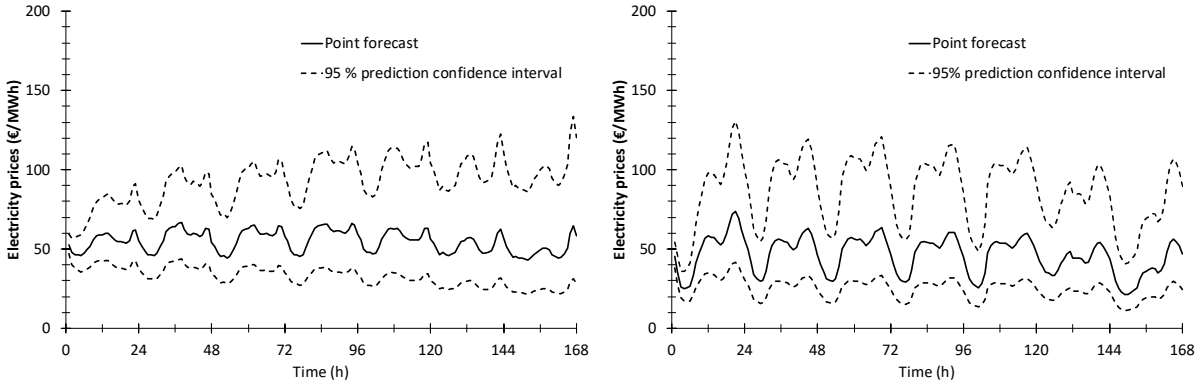
The VPP optimization problem has two uncertain parameters across one week: the electricity price profile and the wind power profile. For each case study, we consider two weeks of interest: Week 1 that refers to the week of August 25-31, 2014; and Week 2 that refers to the week of November 14-30, 2014. Appendix Appendix B describes the sources and sampling process to generate the uncertainty sets and samples (see online supplemental material). These data are available online in [59].

In Figure 2, we provide the point forecast and bounds of the uncertainty set of the electricity prices for the two weeks, and in Figure 3, we show two samples: one with 10 elements and another with 100 elements. Regarding the wind power, we present the original wind power ensemble and the uncertainty set elements In Figure 4, and in Figure 5, we show two samples: one with 10 elements and another with 100 elements.

5.2. Summary of case studies and risk parameterization

In the computational experiments regarding Case 1 and Case 2, we pursue to capture the influence of the technical characteristics of the VPP, and with Week 1 and Week 2, the impact of the wind power and electricity price uncertainty on the performance of the algorithms and the solutions. Overall, we solve two VPP cases, each with two instances of the uncertain parameters.

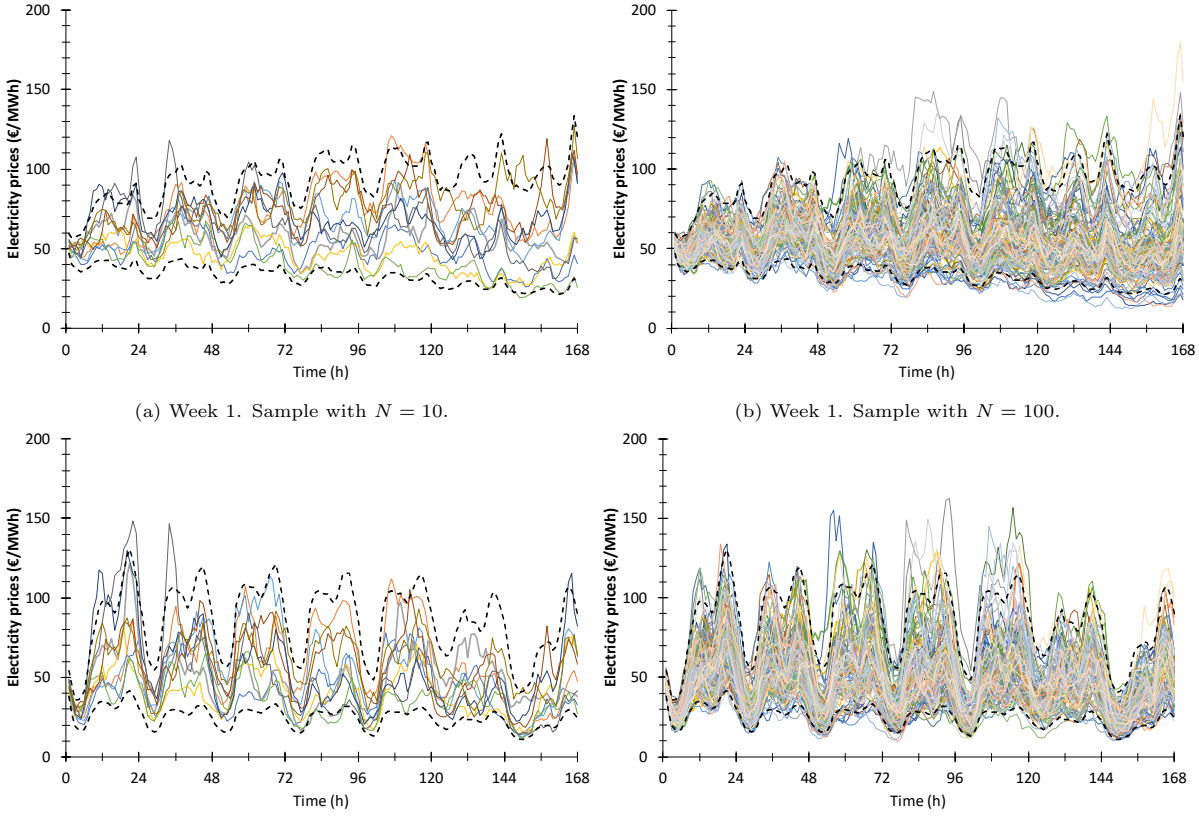
Regarding the SP formulations presented in Section 3.1, they are parameterized over β and α, whereby we consider the cases $\beta = 0$; $\beta = 0.5$ and $\alpha = 0.9$; and $\beta = 1$ and $\alpha = \{0.9; 0.95; 1\}$. The ARO formulation is parameterized over the budget of uncertainty, namely $\Gamma = \{0; 10; 50; 100; 150; 168\}$. Note that there is an additional crucial parameter in the definition of the uncertainty set of the electricity prices: the forecast prediction error level used to define the bounds of the set, which is set to 95%.



(a) Week 1.

(b) Week 2.

Figure 2: Electricity prices point forecast and prediction confidence interval for the two weeks considered.



(a) Week 1. Sample with $N = 10$.

(b) Week 1. Sample with $N = 100$.

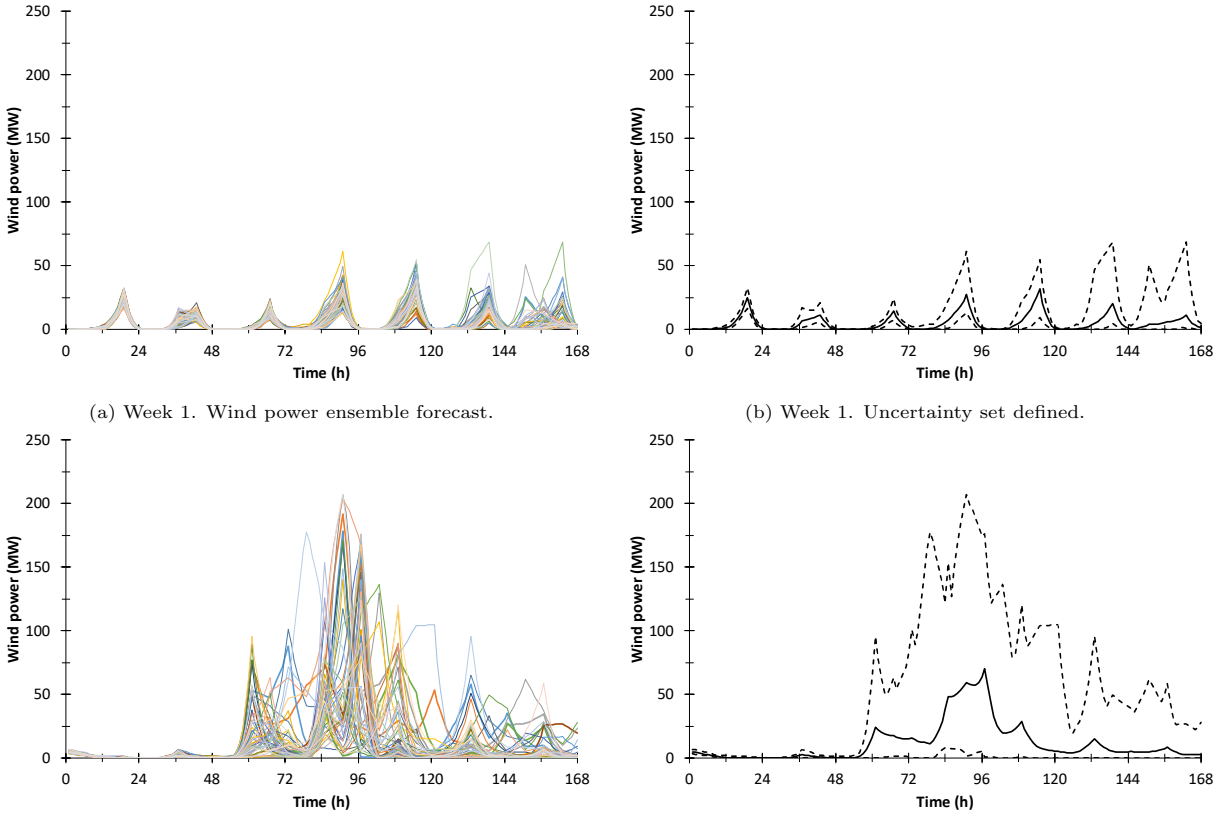
(c) Week 2. Sample with $N = 10$.

(d) Week 2. Sample with $N = 100$.

Figure 3: Examples of two samples for Week 1 and Week 2 used in the SP approach. The sashed lines are the bounds of the uncertainty sets in the ARO approach.

5.3. SAA methodology setup

The results for SP are obtained through an SAA methodology using sample sizes N =10, 50, 100, 500, and 5000. For each sample size, $M = 30$ optimization replications are performed. Each optimization replication involves Algorithm 1 with a different sample. The lower bound on the true optimal objective function value is estimated in two steps: 1) for each distinct first-stage solution obtained from the optimization replications,



(a) Week 1. Wind power ensemble forecast.

(b) Week 1. Uncertainty set defined.

(c) Week 2. Wind power ensemble forecast.

(d) Week 2. Uncertainty set defined.

Figure 4: Wind power ensemble and uncertainty set defined by the average forecast and lower and upper bounds.

a lower bound is estimated using $T = 30$ samples of size $N' = 25,000$, and 2) the first-stage solution with the best lower bound from the previous step is selected and a new lower bound is estimated using $T' = 30$ samples of size $N' = 25,000$. In the second step, estimates of the expected profit, and the CVaR of the profit are also calculated. These samples are available online in [59].

In the ARO approach, the lower bound estimation stage requires only the second step, where estimates of the expected profit and the CVaR of the profit associated with the first-stage variables are calculated.

5.4. Algorithms 1 and 2 setup

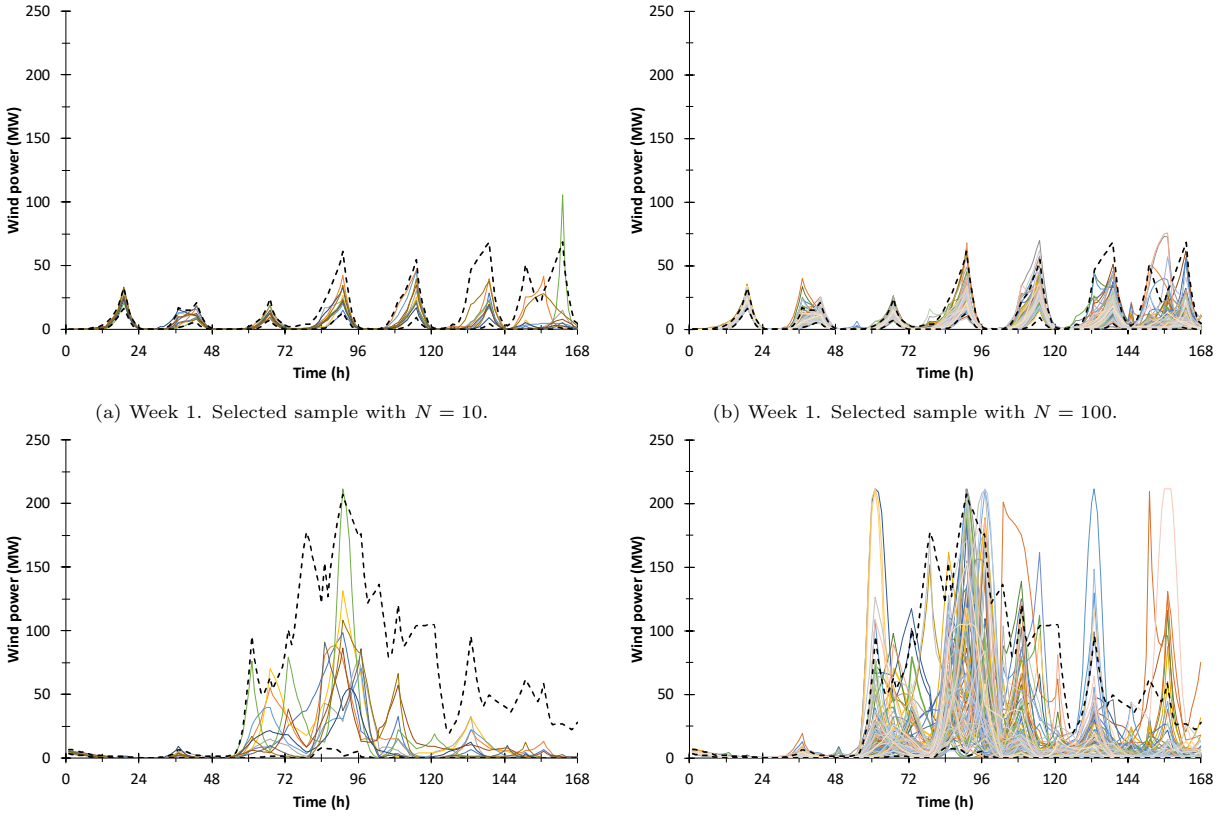
In Algorithms 1 and 2 the stopping criteria are a maximum wall-clock time of 10,800 s, a maximum gap between the bounds of $10^{-4}\%$, and a maximum number of iterations of 5000.

The models and the algorithms were implemented in GAMS, and CPLEX 12.9.0 and GAMS/GRID/GUSS [60] capabilities are used to distribute and solve in parallel the LP subproblems. All optimization runs were performed in the KAUST Ibex computer cluster using exclusive nodes, each with 40 processors Intel Gold 6148 @ 2.6 GHz and 384 Gb of RAM.

5.5. Size of the problems

We provide statistics regarding the size of the problems involved in each decomposition algorithm for different risk parameters values and sample sizes in Appendix Appendix G (online supplemental material).

The size of the master problem increases with the number of iterations, while the size of the subproblems is constant in both approaches. In the first iteration of the decomposition algorithms, the size of the master problem is the same in SP and ARO. However, in SP, depending on the sample size and risk parameters



(a) Week 1. Selected sample with $N = 10$.

(b) Week 1. Selected sample with $N = 100$.

(c) Week 2. Selected sample with $N = 10$.

(d) Week 2. Selected sample with $N = 100$.

Figure 5: Wind power samples generated from the KLE. The dashed lines are the bounds of the uncertainty set.

values, at each iteration, the size of the master problem increases at a different rate. When the CVaR of the profit is considered in the objective function ($\beta > 0$), the number of new constraints and variables added to the master problem at each iteration depends on the sample size. In contrast, for $\beta = 0$ there is only one new constraint added, but with a number of terms that depends on the sample size. In ARO, the size of the master problem increases by the number of primal constraints and variables added on each iteration.

5.6. Comparison of the SP and ARO performance

Detailed results and a discussion on the performance of the ARO decomposition and SP decomposition algorithms are presented in Appendix Appendix E (online supplemental material). Overall, the results show that the relative performance of the SP and ARO models depends on the deterministic and uncertain parameters, sample size, and risk parameters. With all these components, the comparison is not straightforward, and it is challenging to obtain clear-cut conclusions. For simplicity, we focus first on the SP and ARO performance without considering the worst-case in both approaches ($\beta = 1$, $\alpha = 1$ and $\Gamma = 168$), in subsections 5.6.1 and 5.6.2. Then, we analyze specific worst-case results.

5.6.1. Week 1.

In Case 1, the sample size $N = 500$ is a frontier for SP. All sample sizes $N \leq 500$ present lower computational times than ARO; see Figure S6a (online supplemental material). In Case 2, ARO with $\Gamma = \{50, 100, 150\}$ does not converge, and thus requires higher computational times than all SP instances; see Figure S6b (online supplemental material). Also, SP does not meet the gap stopping criterion for one replication out of 30 with $N = 10$, $\beta = 1$, and $\alpha = 0.90$. However, for both Cases, SP with $N = 500$ meets

the gap stopping criterion in all optimization replications of the five sets of risk parameters; see full results in Appendix Appendix H (online supplemental material).

5.6.2. Week 2.

In Case 1 and for sample sizes $N \leq 500$, the computational times for SP are below 10 s, which are lower than the times for the ARO with $\Gamma = \{50, 100\}$ and competitive with $\Gamma = \{10, 150\}$. For Case 2, the SP fails to meet the gap stopping criterion in all sample sizes and risk parameters; see Figure S7b (online supplemental material). The results show that for $\Gamma = \{50, 100\}$, the ARO does not meet the gap stopping criterion, but it does for $\Gamma = \{10, 150, 168\}$; see Table S11 (online supplemental material).

5.7. Comparison between risk-averse SP and ARO.

In this section, we contrast the maximization of the CVaR of the profit with maximization of the worst profit in ARO. The comparison considers different risk levels for the SP: $\alpha = \{0.90, 0.95, 1\}$; and the ARO: $\Gamma = \{10, 50, 100, 150, 168\}$.

Figures 6 and 7 show the results for Cases 1 and 2 and Weeks 1 and 2 for $M = 30$ SP optimization replications with $N = \{10, 50, 100, 500, 5000\}$. For each case, there are three subfigures for $\alpha = \{0.90, 0.95, 1\}$.

5.7.1. Week 1.

For Case 1, Figures 6a and 6c show that for $\alpha = 0.90$ and $\alpha = 0.95$ only the SP computational times of $N = 5000$ are worse than the ARO ones. For $\alpha = 1$ the variance of the computational times with $N = 500$ and $N = 5000$ increases considerably; see Figure 6e. However, for $N = \{10, 50, 100\}$, most of the computational times are below 10 s, which makes it competitive with the worst-case in ARO, $\Gamma = 168$.

For Case 2, the variance of the computational times for the SP is larger than for Case 1. This variance is clear for $\alpha = 0.90$ for $N = 10$; see Figure 6b, and for all sample sizes for $\alpha \geq 0.95$. For $\alpha = 1$, all sample sizes lead to significant variances between any two optimization replications; see Figure 6f. The ARO with $\Gamma = \{50, 100, 150, 168\}$ performs worse than SP, but the SP for $\alpha = 1$ leads to a non-predictable computational time, which means that one simple optimization replication is not sufficient to compare with the ARO computational time.

These results for Case 2, Week 1 show a link between the performance of SP and ARO when the worst-case, or near worst-case, is considered. In both approaches, the computational times increase considerably. This increment in the time as the risk-averse parameters are selected is clear in Table S9 (online supplemental material). This contrasts with the risk-neutral SP, where the objective function accounts for the expected profit, and therefore, the lower electricity prices and search for the worst profit do not drive the optimization.

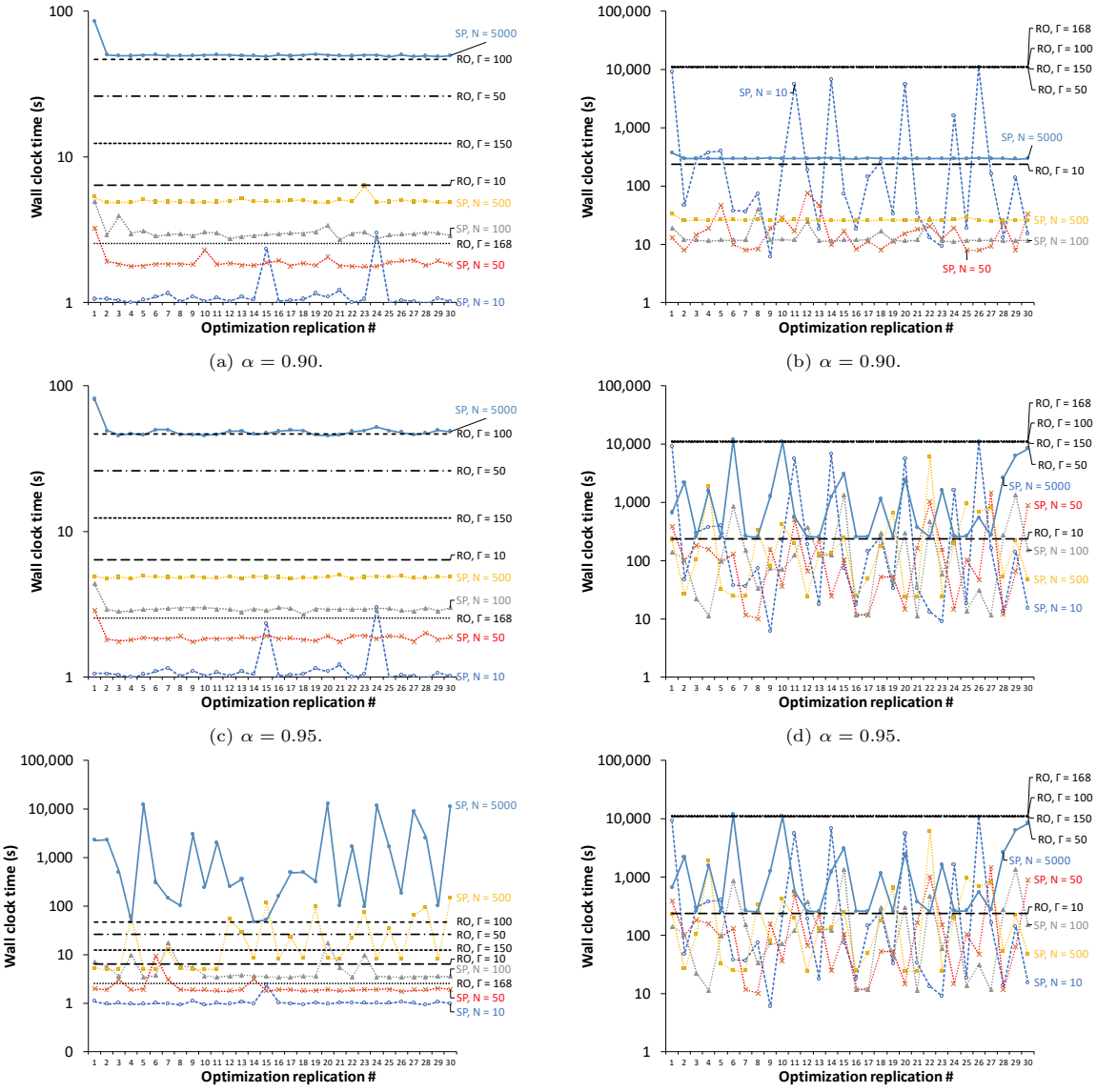
5.7.2. Week 2.

The trend of the results in Case 1 is similar to Week 1. In Case 2, the ARO meets the gap stopping criterion with $\Gamma = \{10, 150, 168\}$, whereas, SP did not in most replications. The exceptions are some replications with $N = 10$ and $N = 50$. Overall, the final gaps of SP are over 22% for $\beta = 1$, $\alpha = 1$; see Table S11 (online supplemental material).

5.8. Risk management analysis

The VPP problem is formulated as a two-stage decision framework, where the relevant part of the optimal solutions is the first-stage solution. The first-stage solution is implemented before the planning horizon, whereas the second-stage solution is just the optimal recourse solution for each element of the sample in SP and the worst realization in the ARO. Therefore, we limit our discussion to the optimal first-stage solutions and the estimates of the expected profit and CVaR of the profit.

In Figures 8 and 9, we present the estimates of the expected profit and CVaR of the profit obtained from the bound estimation stage of the SAA methodology, for the first-stage solutions presented in Tables 2, 3, 4, and 5. In these figures, we illustrate with arrows the price of robustness and the value of robustness obtained with SP and ARO for multiple risk-aversion levels. The concept of "price of robustness" was



(a) $\alpha = 0.90$.

(b) $\alpha = 0.90$.

(c) $\alpha = 0.95$.

(d) $\alpha = 0.95$.

(e) $\alpha = 1$.

(f) $\alpha = 1$.

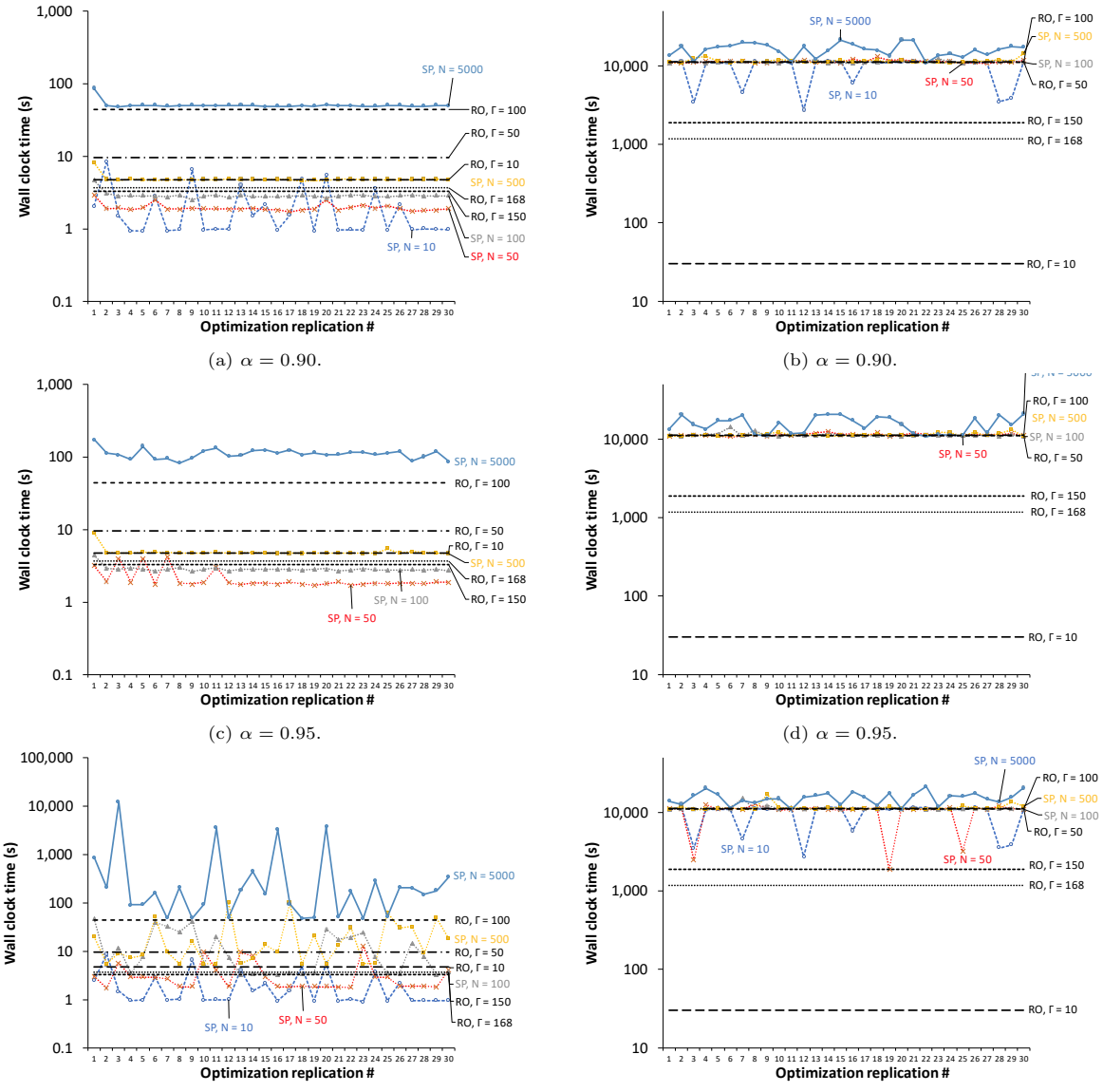
Figure 6: Left: Case 1, Week 1; Right: Case 2, Week 1. Wall clock time for $M = 30$ risk-averse SP optimization replications with $\beta = 1$, and ARO.

introduced by Bertsimas et al. [47] to measure the deviation between the objective function solutions from a robust optimization problem and the corresponding deterministic problem. Similarly, Gregory et al. [61] proposed two measures for the "cost of robustness" that measure the deviation between a non-robust solution and a robust one, applied to portfolio return optimization problems. We define the price of robustness as the deviation between the estimates of the expected profit for $\beta = 0$ and $\beta > 0$ for SP:

$$\text{Price of robustness}(\alpha, \beta) := \mathbb{E}[\text{Profit}]\big|_{\beta_{\text{fixed}}=0} - \mathbb{E}[\text{Profit}]\big|_{\alpha,\beta>0}, \tag{22}$$

and between $\Gamma = 0$ and $\Gamma > 0$ for ARO:

$$\text{Price of robustness}(\Gamma) := \mathbb{E}[\text{Profit}]\big|_{\Gamma_{\text{fixed}}=0} - \mathbb{E}[\text{Profit}]\big|_{\Gamma>0}. \tag{23}$$



(a) $\alpha = 0.90$. (b) $\alpha = 0.90$.

(c) $\alpha = 0.95$. (d) $\alpha = 0.95$.

(e) $\alpha = 1$. (f) $\alpha = 1$.

Figure 7: Left: Case 1, Week 2; Right: Case 2, Week 2. Wall clock time for $M = 30$ risk-averse SP optimization replications with $\beta = 1$, and ARO.

We introduce the value of robustness to measure the deviation between the estimates of the worst profit for $\beta > 0$ and $\beta = 0$ for SP:

$$\text{Value of robustness}(\alpha, \beta) := \text{CVaR}_{\alpha=1}[\text{Profit}]\big|_{\alpha,\beta>0} - \text{CVaR}_{\alpha'=1}[\text{Profit}]\big|_{\beta_{\text{fixed}}=0}, \tag{24}$$

and between $\Gamma = 0$ and $\Gamma > 0$ for ARO:

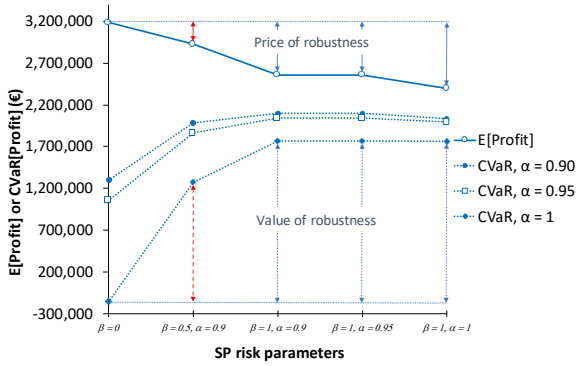
$$\text{Value of robustness}(\Gamma) := \text{CVaR}_{\alpha'=1}[\text{Profit}]\big|_{\Gamma>0} - \text{CVaR}_{\alpha=1}[\text{Profit}]\big|_{\Gamma_{\text{fixed}}=0}. \tag{25}$$

In both SP and ARO, the value of robustness is higher than the price of robustness for each set of risk parameters. For example, for Case 1, Week 1, we obtain an expected price of robustness of 264,052 € and a value of robustness of 1,436,451 €, with the solution from the SP model with $\beta = 0.5$, $\alpha = 0.9$; see Figure

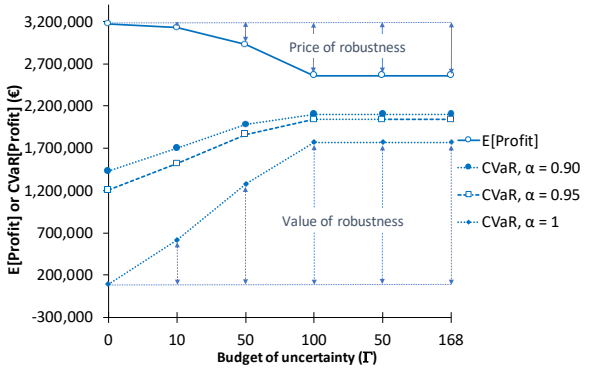
8a. These values result from the first-stage decisions to buy electricity through forward contracts (see Table 2) to sell it in the pool (or used by the hydro-pumped storage unit and then sold at a lower electricity price).

Comparing SP and ARO results, the profiles of expected profit and CVaR of profit have a similar trend and some values in common. In Case 1, the profiles are similar because of a relevant insight: the SP and ARO lead to the same first-stage solutions, which means that the same solution can be obtained in SP and ARO depending on the risk parameterization. The the first-stage solutions are presented in Tables 2 and 3 for Weeks 1 and 2, respectively. In these tables, we highlight similar solutions obtained with SP and ARO, which lead to the same estimates for the expected profit and CVaR of the profit in the bound estimation stage. In Case 2, Week 1, the first-stage solutions UT, SU, SD, and SC of the SP and ARO are similar, but the variable BC is different; this leads to distinct estimates of the expected profit and CVaR of the profit; see Table 4. In Case 2, Week 2, there are differences between the first-stage solutions of SP and ARO; see Table 5. However, definite conclusions cannot be drawn because the decomposition algorithms did not converge to small gaps. As the robustness increases, the confidence intervals become tighter for both estimates: expectation and CVaR; see the column labeled $\pm$ in Tables 2 to 5. This reduction in the interval occurs because as the robustness increases, the electricity to sell by contract increases, and therefore, it reduces the uncertainty of the profit; see the column labeled SC in those tables.

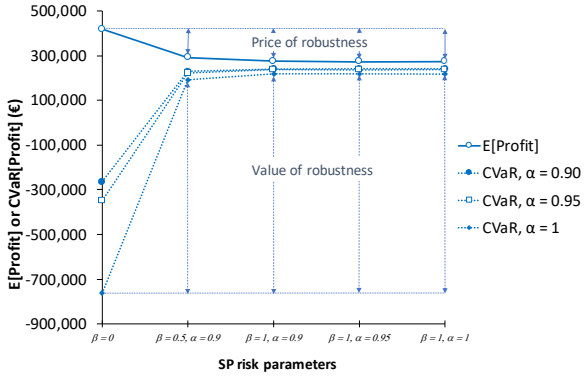
The average, median, and standard deviation of the wall clock time required to calculate the estimates and confidence intervals for one first-stage solution are 1329, 656, and 2315 seconds, respectively, based on the wall clock times for the second step described in Section 5.3. The wall clock time for one first stage solution depends mainly on the samples' sizes and number of replications, which were $N' = 25,000$, and



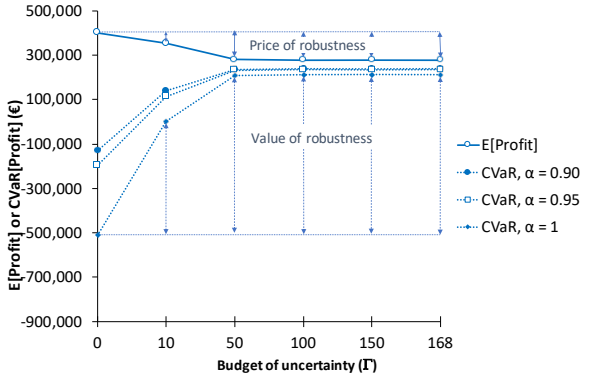
(a) **SP**, Case 1, Week 1. Red arrows: price of robustness of 264,052 € and a value of robustness of 1,436,451 €.



(b) **ARO**, Case1, Week 1.

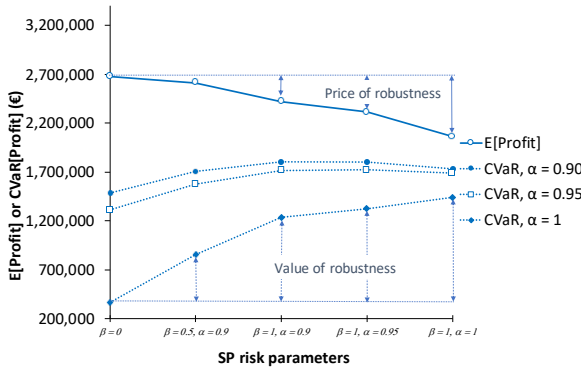


(c) **SP**, Case 2, Week 1.

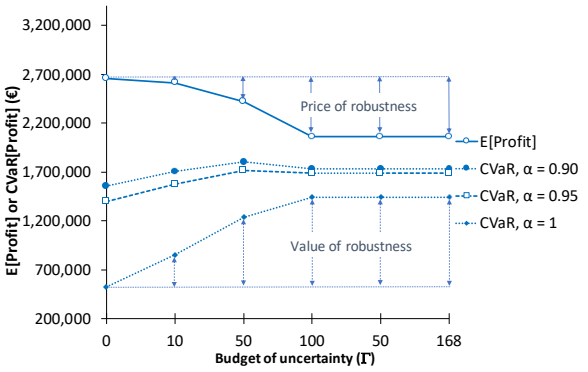


(d) **ARO**, Case 2, Week 1.

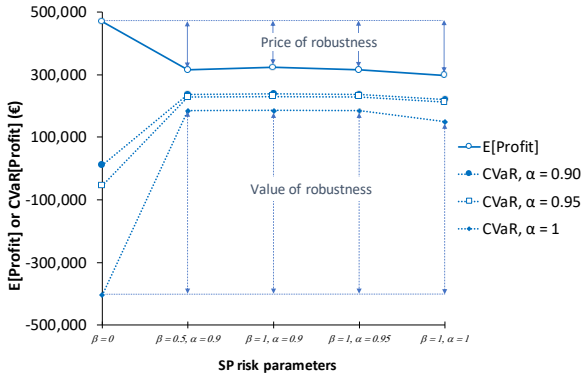
Figure 8: Expected profit and CVaR of the profit as a function of the formulation and risk parameters. $N = 500$, $M = 30$, $\{T, T'\} = 30$, $N' = 25,000$.



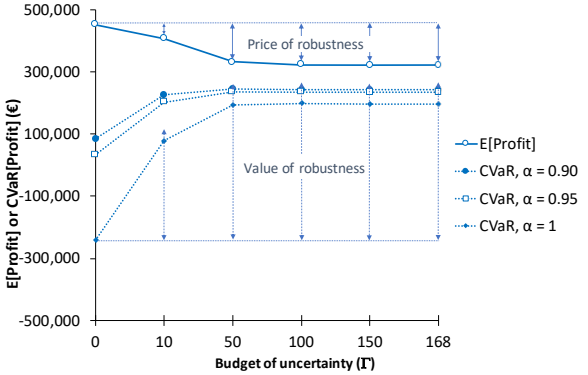
(a) **SP**, Case 1, Week 2.



(b) **ARO**, Case1, Week 2.



(c) **SP**, Case 2, Week 2.



(d) **ARO**, Case 2, Week 2.

Figure 9: Expected profit and CVaR of the profit as a function of the formulation and risk parameters. $N = 500$, $M = 30$, $\{T, T'\} = 30$, $N' = 25,000$.

$T' = 30$. That is the time to solve 30 times 25,000 LP problems, plus the time to calculate the estimates for the expectation and CVaR. A comprehensive study on these computational times was reported in Lima et al. [55]. We provide additional results obtained with the SAA methodology in Appendices Appendix I and Appendix J (online supplemental material). Specifically, on the quality of the SP solutions using confidence intervals and bounds on w^*, and results from all optimization replications with $N = 500$.

Table 2: Case 1, Week 1. First-stage solution, estimates, and confidence intervals. Similar solutions in SP and ARO are in **bold**. $N' = 25,000$, $\{T, T'\} = 30$.

							Estimates and confidence intervals					
SP		First-stage solution							CVaR[Profit] (€)			
β	α	UT	SU	SD	SC	BC	$\mathbb{E}$[Profit] (€)	$\pm$	$\alpha = 0.95$	$\pm$	$\alpha = 1$	$\pm$
0	-	100	0	0	0	160	3,188,599	2730	1,055,341	3875	-159,189	43,486
0.5	0.9	**100**	**0**	**0**	**155**	**0**	**2,924,547**	**1365**	**1,865,146**	**1922**	**1,277,262**	**20,447**
1	0.9	**100**	**0**	**0**	**315**	**0**	**2,557,799**	**672**	**2,043,343**	**934**	**1,768,962**	**8995**
1	0.95	**100**	**0**	**0**	**315**	**0**	**2,557,799**	**672**	**2,043,343**	**934**	**1,768,962**	**8995**
1	1	95.24	1	1	315	0	2,397,745	510	1,993,444	790	1,764,836	9268
	RO											
Γ												
0	-	100	0	0	0	110	3,170,852	2513	1,208,071	3565	93,103	39,771
10	-	100	0	0	50	55	3,124,555	2058	1,519,757	2914	613,889	32,066
50	-	**100**	**0**	**0**	**155**	**0**	**2,924,547**	**1365**	**1,865,146**	**1,922**	**1,277,262**	**20,447**
100	-	**100**	**0**	**0**	**315**	**0**	**2,557,799**	**672**	**2,043,343**	**934**	**1,768,962**	**8995**
150	-	**100**	**0**	**0**	**315**	**0**	**2,557,799**	**672**	**2,043,343**	**934**	**1,768,962**	**8995**
168	-	**100**	**0**	**0**	**315**	**0**	**2,557,799**	**672**	**2,043,343**	**934**	**1,768,962**	**8995**

UT (%) - up-time of the thermal unit; SU - number of startups of the thermal unit; SD - number of shutdowns of the thermal unit; SC (MW) - power sold through contract; BC (MW) - power bought through contract.

Table 3: Case 1, Week 2. First-stage solution, estimates, and confidence intervals. Similar solutions in SP and ARO are in **bold**. $N' = 25,000$, $\{T, T'\} = 30$.

							Estimates and confidence intervals					
SP		First-stage solution							CVaR[Profit] (€)			
β	α	UT	SU	SD	SC	BC	$\mathbb{E}$[Profit] (€)	$\pm$	$\alpha = 0.95$	$\pm$	$\alpha = 1$	$\pm$
0	-	100	0	0	0	160	2,675,583	1,658	1,312,720	2,986	366,264	41,289
0.5	0.9	**100**	**0**	**0**	**50**	**55**	**2,611,946**	**1,265**	**1,573,094**	**2,278**	**855,899**	**31,207**
1	0.9	**100**	**0**	**0**	**155**	**0**	**2,420,072**	**862**	**1,714,915**	**1,548**	**1,234,206**	**20,671**
1	0.95	100	0	0	210	0	2,313,649	724	1,722,800	1,298	1,323,457	17,216
1	1	**100**	**0**	**0**	**315**	**0**	**2,061,459**	**464**	**1,687,431**	**837**	**1,440,562**	**10,633**
	RO											
Γ												
0	-	100	0	0	0	110	2,655,396	1,531	1,397,100	2,758	524,552	38,019
10	-	**100**	**0**	**0**	**50**	**55**	**2,611,946**	**1,265**	**1,573,094**	**2,278**	**855,899**	**31,207**
50	-	**100**	**0**	**0**	**155**	**0**	**2,420,072**	**862**	**1,714,915**	**1,548**	**1,234,206**	**20,671**
100	-	**100**	**0**	**0**	**315**	**0**	**2,061,459**	**464**	**1,687,431**	**837**	**1,440,562**	**10,633**
150	-	**100**	**0**	**0**	**315**	**0**	**2,061,459**	**464**	**1,687,431**	**837**	**1,440,562**	**10,633**
168	-	**100**	**0**	**0**	**315**	**0**	**2,061,459**	**464**	**1,687,431**	**837**	**1,440,562**	**10,633**

UT (%) - up-time of the thermal unit; SU - number of startups of the thermal unit; SD - number of shutdowns of the thermal unit; SC (MW) - power sold through contract; BC (MW) - power bought through contract.

Table 4: Case 2, Week 1. First-stage solution, estimates, and confidence intervals. $N' = 25,000$, $\{T, T'\} = 30$.

SP		First-stage solution					Estimates and confidence intervals					
									CVaR[Profit] (€)			
β	α	UT	SU	SD	SC	BC	$\mathbb{E}$[Profit] (€)	$\pm$	$\alpha = 0.95$	$\pm$	$\alpha = 1$	$\pm$
0	-	100	1	0	0	160.0	418,652	996	-348,002	1370	-761,885	15,169
0.5	0.9	100	1	0	100	52.0	291,036	108	222,394	129	190,888	1320
1	0.9	100	1	0	100	38.6	274,003	65	236,094	74	217,293	733
1	0.95	100	1	0	100	38.3	273,523	64	236,112	72	217,366	753
1	1	100	1	0	100	38.0	273,237	64	236,106	71	217,387	769
RO												
Γ												
0	-	100	1	0	0	110.0	400,905	780	-195,529	1,062	-510,493	11,416
10	-	100	1	0	50	55.0	354,608	327	114,289	423	1,790	3,820
50	-	100	1	0	100	44.7	281,773	82	232,175	98	209,002	981
100	-	100	1	0	100	42.7	279,164	76	234,112	90	212,975	886
150	-	100	1	0	100	42.2	278,492	75	234,523	88	213,881	867
168	-	100	1	0	100	42.2	278,503	75	234,517	88	213,866	868

UT (%) - up-time of the thermal unit; SU - number of startups of the thermal unit; SD - number of shutdowns of the thermal unit; SC (MW) - power sold through contract; BC (MW) - power bought through contract.

Table 5: Case 2, Week 2. First-stage solution, estimates, and confidence intervals. $N' = 25,000$, $\{T, T'\} = 30$.

SP		First-stage solution					Estimates and confidence intervals					
									CVaR[Profit] (€)			
β	α	UT	SU	SD	SC	BC	$\mathbb{E}$[Profit] (€)	$\pm$	$\alpha = 0.95$	$\pm$	$\alpha = 1$	$\pm$
0	-	97.02	3	2	0	160	469,302	646	-54,611	1,151	-404,614	14,623
0.5	0.9	83.33	10	9	57.2	0	314,255	141	227,195	144	184,731	1,969
1	0.9	80.95	10	9	50.0	0	322,659	148	229,509	158	186,025	1,881
1	0.95	80.95	8	7	59.1	0	314,158	142	228,916	134	185,681	1,897
1	1	76.19	8	7	70.1	0	297,171	132	212,157	198	150,244	3,003
RO												
Γ												
0	-	89.88	2	1	0	110	451,526	518	33,832	920	-240,902	11,595
10	-	89.88	2	1	50	55	408,076	269	202,229	443	78,897	5,089
50	-	97.62	1	0	58.1	0	332,841	157	235,824	172	192,975	1,898
100	-	98.21	1	0	65.6	0	323,572	148	235,754	136	198,186	1,573
150	-	100	1	0	66.0	0	321,970	147	234,460	135	197,047	1,571
168	-	100	1	0	66.0	0	321,986	147	234,461	135	197,043	1,571

UT (%) - up-time of the thermal unit; SU - number of startups of the thermal unit; SD - number of shutdowns of the thermal unit; SC (MW) - power sold through contract; BC (MW) - power bought through contract.

6. Conclusions

This work presented a comparison between two-stage risk-averse SP and two-stage ARO applied to the optimal scheduling problem of a VPP. We conducted an extensive analysis. We presented in detail the formulations used in both approaches and discussed algorithmic details, including the type of cuts and size of the master problems in the decomposition algorithms. The decomposition algorithm implemented for the SP uses an efficient parallel solution of the LP subproblems, whereas the ARO decomposition algorithm solves an MILP subproblem. Fundamentally, they both rely on solving LP problems at the subproblem level.

Regarding the analysis of the computational performance trends, we covered different parameterizations of the risk in the SP approach, distinct values for the budget of uncertainty parameter in the ARO approach, and multiple samples and sample sizes in the SP.

The relative performance of the two approaches falls in three cases. In the first case, both approaches are competitive. If the sample size $N \leq 500$, then the SP model is in general faster. This indicates that if both approaches solve the problem to meet the gap criterion, then the relative performance of SP and ARO depends on the efficiency to solve N LP subproblems in parallel in SP vs. solving the MILP subproblem in ARO. However, the extreme SP risk-averse formulation is less efficient, and requires a smaller sample size to be comparable to ARO. This occurs in Case 1, Week 1 and Week 2. In the second case, the SP approach is comparatively less effective than in the first case, but it still converges in most of the replications, whereas the ARO does not meet the gap stopping criterion in four out of five sets of risk parameters. The performance of the extreme SP risk-averse formulations is inferior to those of the other SP formulations, which is aligned with the difficulties of the ARO. This case occurs in Case 2, Week 1. In the third case, the SP approach fails to meet the gap stopping criterion independently of the sample size, whereas the ARO meets the gap stopping criterion in three out of five sets of risk parameters. Here, even using a sample size $N = 10$, SP cannot meet the gap stopping criterion. Therefore, this SP performance is an indicator of ARO superiority in this case. This case occurs in Case 2, Week 2. In general, if the problem is amenable to be solved by both approaches, then the SP is generally more efficient below a sample size threshold. On the other hand, if the SP formulation is difficult to solve with smaller sample sizes, ARO is generally more efficient. These computational results suggest that the relative performance of both approaches depends on: a) the combined effect of the VPP data and the uncertainty data; and b) the risk parameter values.

A relevant conclusion is that similar results were obtained with the SP and ARO approaches. The results show that by adjusting the risk parameters in both approaches, it is possible to obtain similar first-stage optimal solutions and estimates of the expected value and CVaR of the profit. This conclusion is evident in Case 1. For Case 2, there is a subset of variables of the first-stage solution that is similar in the SP and ARO approaches, but there is a difference in one variable of the first-stage solution. However, a clear-cut conclusion cannot be drawn when the decomposition algorithms did not meet the gap stopping criterion.

The results indicate that both approaches are suitable to determine risk-averse solutions. Overall, the choice of one approach over the other should first be based on the selected representation of the uncertainty, given that both approaches can be competitive depending on the problem data (deterministic and uncertain), sample size, and risk-aversion level.

Acknowledgements

Research reported in this publication was supported by research funding from KAUST and computational resources from the Supercomputing Laboratory from KAUST Core Labs. Figure 1 was produced by Xavier Pita, scientific illustrator at KAUST.

References

[1] S. Awerbuch, A. Preston (Eds.), The Virtual Utility: Accounting, Technology & Competitive Aspects of the Emerging Industry, Springer US, 1997.

[2] D. Pudjianto, C. Ramsay, G. Strbac, Virtual power plant and system integration of distributed energy resources, IET Renew. Power Gener. 1 (2007) 10–16.

[3] K. Dietrich, J. M. Latorre, L. Olmos, A. Ramos, Modelling and assessing the impacts of self supply and market-revenue driven virtual power plants, Electric Power Systems Research 119 (2015) 462 – 470.

[4] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, M. Zugno, Integrating Renewables in Electricity Markets, Springer US, 2014.

[5] W. Zhang, H. Rahimian, G. Bayraksan, Decomposition Algorithms for Risk-Averse Multistage Stochastic Programs with Application to Water Allocation under Uncertainty, INFORMS J. Comput. 28 (2016) 385–404.

[6] R. M. Lima, A. J. Conejo, S. Langodan, I. Hoteit, O. M. Knio, Risk-averse formulations and methods for a virtual power plant, Comput. Oper. Res. 96 (2018) 350 – 373.

[7] D. Shang, V. Kuzmenko, S. Uryasev, Cash flow matching with risks controlled by buffered probability of exceedance and conditional value-at-risk, Ann. Oper. Res. 260 (2018) 501–514.

[8] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, T. Zheng, Adaptive Robust Optimization for the Security Constrained Unit Commitment Problem, IEEE T. Power Syst. 28 (2013) 52–63.

[9] Q. Zhang, I. E. Grossmann, R. M. Lima, On the Relation Between Flexibility Analysis and Robust Optimization for Linear Systems, AIChE J. 62 (2016) 3109–3123.

[10] B. L. Gorissen, I. Yanikoglu, D. den Hertog, A practical guide to robust optimization, Omega-Int. J. Manage. S. 53 (2015) 124–137.

[11] V. Gabrel, C. Murat, A. Thiele, Recent advances in robust optimization: An overview, Eur. J. Oper. Res. 235 (2014) 471–483.

[12] A. Ben-Tal, L. El Ghaoui, A. Nemirovski, Robust Optimization, Princeton Series in Applied Mathematics, Princeton University Press, 2009.

[13] H. Pandžić, J. M. Morales, A. J. Conejo, I. Kuzle, Offering model for a virtual power plant based on stochastic programming, Appl. Energ. 105 (2013) 282–292.

[14] M. A. Tajeddini, A. Rahimi-Kian, A. Soroudi, Risk averse optimal operation of a virtual power plant using two stage stochastic programming, Energy 73 (2014) 958–967.

[15] J. Z. Riveros, K. Bruninx, K. Poncelet, W. D'haeseleer, Bidding strategies for virtual power plants considering CHPs and intermittent renewables, Energ. Convers. Manage. 103 (2015) 408–418.

[16] E. G. Kardakos, C. K. Simoglou, A. G. Bakirtzis, Optimal Offering Strategy of a Virtual Power Plant: A Stochastic Bi-Level Approach, IEEE T. Smart Grid 7 (2016) 794–806.

[17] S. R. Dabbagh, M. K. Sheikh-El-Eslami, Risk Assessment of Virtual Power Plants Offering in Energy and Reserve Markets, IEEE T. Power Syst. 31 (2016) 3572–3582.

[18] A. Castillo, J. Flicker, C. W. Hansen, J.-P. Watson, J. Johson, Stochastic optimization with risk aversion for virtual power plant operations: A rolling horizon control, IET Gener. Transm. Dis. (2018). Accepted, available online: 27 November 2018.

[19] R. M. Lima, A. Q. Novais, A. J. Conejo, Weekly self-scheduling, forward contracting, and pool involvement for an electricity producer. An adaptive robust optimization approach, Eur. J. Oper. Res. 240 (2015) 457–475.

[20] M. Shabanzadeh, M.-K. Sheikh-El-Eslami, M.-R. Haghifam, The design of a risk-hedging tool for virtual power plants via robust optimization approach, Appl. Energ. 155 (2015) 766–777.

[21] M. Rahimiyan, L. Baringo, Strategic Bidding for a Virtual Power Plant in the Day-Ahead and Real-Time Markets: A Price-Taker Robust Optimization Approach, IEEE T. Power Syst. 31 (2016) 2676–2687.

[22] A. Baringo, L. Baringo, A Stochastic Adaptive Robust Optimization Approach for the Offering Strategy of a Virtual Power Plant, IEEE T. Power Syst. 32 (2017) 3492–3504.

[23] X. Zhang, A. J. Conejo, Robust Transmission Expansion Planning Representing Long-and Short-Term Uncertainty, IEEE T. Power Syst. 33 (2018) 1329–1338.

[24] B. Fanzeres, A. Street, L. A. Barroso, Contracting Strategies for Renewable Generators: A Hybrid Stochastic and Robust Optimization Approach, IEEE T. Power Syst. 30 (2015) 1825–1837.

[25] C. Zhao, Y. Guan, Unified Stochastic and Robust Unit Commitment, IEEE T. Power Syst. 28 (2013) 3353–3361.

[26] G. Lagos, D. Espinoza, E. Moreno, J. Amaya, Robust planning for an open-pit mining problem under ore-grade uncertainty, Electronic Notes in Discrete Mathematics 37 (2011) 15 – 20. LAGOS'11 – VI Latin-American Algorithms, Graphs and Optimization Symposium.

[27] H. Pandzic, Y. Dvorkin, T. Qiu, Y. Wang, D. S. Kirschen, Toward Cost-Efficient and Reliable Unit Commitment Under Uncertainty, IEEE T. Power Syst. 31 (2016) 970–982.

[28] W. van Ackooij, A comparison of four approaches from stochastic programming for large-scale unit-commitment, EURO J. Comput. Optim. 5 (2017) 119–147.

[29] N. Kazemzadeh, S. Ryan, M. Hamzeei, Robust optimization vs. stochastic programming incorporating risk measures for unit commitment with uncertain variable renewable generation, Energy Syst. 10 (2019) 517–541.

[30] J. R. Birge, F. Louveaux, Introduction to stochastic programming, Springer, 2011.

[31] E. Hope, L. Rud, B. Singh, Models for Energy Policy, Routledge, London, 1996, pp. 202–210.

[32] S. Deng, S. Oren, Electricity derivatives and risk management, Energy 31 (2006) 940–953.

[33] T. Kristiansen, Pricing of monthly forward contracts in the Nord Pool market, Energ. Policy 35 (2007) 307–316.

[34] F. E. Benth, S. Koekebakker, Stochastic modeling of financial electricity contracts, Energ. Econ. 30 (2008) 1116–1157.
[35] F. E. Benth, A. Cartea, R. Kiesel, Pricing forward contracts in power markets by the certainty equivalence principle: Explaining the sign of the market risk premium, J. Bank. Financ. 32 (2008) 2006–2021. International Summer School on Risk Measurement and Control, Rome, Italy, Jun, 2007.
[36] A. Botterud, T. Kristiansen, M. D. Ilic, The relationship between spot and futures prices in the Nord Pool electricity market, Energ. Econ. 32 (2010) 967–978.
[37] R. Weron, M. Zator, Revisiting the relationship between spot and futures prices in the Nord Pool electricity market, Energ. Econ. 44 (2014) 178–190.
[38] R. Aïd, Electricity Markets, Springer International Publishing, Cham, 2015, pp. 5–25.
[39] C. Pflug, Probabilistic Constrained Optimization: Methodology and Applications, Springer US, 2000, pp. 272–281.
[40] R. T. Rockafellar, S. Uryasev, Optimization of conditional value-at-risk, J. Risk 2 (2000) 21–41.
[41] R. T. Rockafellar, Coherent Approaches to Risk in Optimization Under Uncertainty, INFORMS, 2007, pp. 38–61. arXiv:http://pubsonline.informs.org/doi/pdf/10.1287/educ.1073.0032.
[42] P. Artzner, F. Delbaen, J. Eber, D. Heath, Coherent measures of risk, Math. Financ. 9 (1999) 203–228.
[43] D. Bertsimas, V. Gupta, N. Kallus, Data-driven robust optimization, Math. Program. 167 (2018) 235–292.
[44] D. Bertsimas, R. Shioda, Algorithm for cardinality-constrained quadratic optimization, Comput. Optim. Appl. 43 (2009) 1–22.
[45] A. Thiele, T. Terry, M. Epelman, Robust linear optimization with recourse, Technical Report, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48019, USA, 2009. Technical Report TR09-01, March 2009.
[46] R. Jiang, J. Wang, Y. Guan, Robust Unit Commitment With Wind Power and Pumped Storage Hydro, IEEE T. Power Syst. 27 (2012) 800–810.
[47] D. Bertsimas, D. Pachamanova, M. Sim, Robust linear optimization under general norms, Oper. Res. Lett. 32 (2004) 510–516.
[48] D. Bertsimas, M. Sim, Tractable approximations to robust conic optimization problems, Math. Program. 107 (2006) 5–36.
[49] M. Zugno, J. M. Morales, H. Madsen, Commitment and dispatch of heat and power units via affinely adjustable robust optimization, Comput. Oper. Res. 75 (2016) 191–201.
[50] S. Zhang, S. Wang, Q. Xu, A New Reactive Scheduling Approach for Short-Term Crude Oil Operations under Tank Malfunction, Ind. Eng. Chem. Res. 54 (2015) 12438–12454.
[51] R. Jiang, M. Zhang, G. Li, Y. Guan, Two-stage network constrained robust unit commitment problem, Eur. J. Oper. Res. 234 (2014) 751–762.
[52] M. Zugno, A. J. Conejo, A robust optimization approach to energy and reserve dispatch in electricity markets, Eur. J. Oper. Res. 247 (2015) 659–671.
[53] A. Lorca, X. A. Sun, Adaptive Robust Optimization With Dynamic Uncertainty Sets for Multi-Period Economic Dispatch Under Significant Wind, IEEE T. Power Syst. 30 (2015) 1702–1713.
[54] C. I. Fábián, Handling CVaR objectives and constraints in two-stage stochastic models, Eur. J. Oper. Res. 191 (2008) 888–911.
[55] R. M. Lima, A. J. Conejo, L. Giraldi, O. Le Maître, I. Hoteit, O. M. Knio, Sample average approximation for risk-averse problems: A virtual power plant scheduling application, EURO J. Comput. Optim. 9 (2021) 100005.
[56] A. Kleywegt, A. Shapiro, T. Homem-De-Mello, The sample average approximation method for stochastic discrete optimization, SIAM J. Optimiz. 12 (2001) 479–502.
[57] R. Van Slyke, R. Wets, L-Shaped Linear Programs with Applications to Optimal Control and Stochastic Programming, SIAM J. Appl. Math. 17 (1969) 638–663.
[58] B. Zeng, L. Zhao, Solving two-stage robust optimization problems using a column-and-constraint generation method, Oper. Res. Lett. 41 (2013) 457–461.
[59] R. M. Lima, A. J. Conejo, L. Giraldi, O. Le Maître, I. Hoteit, O. M. Knio, Samples of wind power and electricity prices used in the manuscript "Risk-averse stochastic programming vs. adaptive robust optimization: a virtual power plant application", King Abdullah University of Science and Technology (KAUST) repository, 2021. URL: http://hdl.handle.net/10754/673934. doi:10.25781/KAUST-3O88T, http://hdl.handle.net/10754/673934.
[60] M. R. Bussieck, M. C. Ferris, T. Lohmann, Algebraic Modeling Systems: Modeling and Solving Real World Optimization Problems, Springer, Berlin Heidelberg, 2012, pp. 35–56.
[61] C. Gregory, K. Darby-Dowman, G. Mitra, Robust optimization and portfolio selection: The cost of robustness, Eur. J. Oper. Res. 212 (2011) 417–428.
[62] A. J. Conejo, R. Garcia-Bertrand, M. Carrion, A. Caballero, A. de Andres, Optimal involvement in futures markets of a power producer, IEEE T. Power Syst. 23 (2008) 703–711.
[63] G. E. P. Box, G. Jenkins, Time Series Analysis, Forecasting and Control, Holden-Day, Incorporated, 1990.
[64] Iberian Electricity Market, www.omie.es, 2015.
[65] A. J. Conejo, M. Carrión, J. M. Morales, Decision Making Under Uncertainty in Electricity Markets, Springer US, 2010.
[66] R. Weron, Electricity price forecasting: A review of the state-of-the-art with a look into the future, Int. J. Forecasting 30 (2014) 1030–1081.
[67] O. Le Maître, O. M. Knio, Spectral methods for uncertainty quantification: with applications to computational fluid dynamics, Springer Science & Business Media, 2010.

Supplemental material

This document presents supplementary material for the manuscript "Risk-averse stochastic programming vs. adaptive robust optimization: a virtual power plant application" by Ricardo M. Lima, Antonio J. Conejo, Loïc Giraldi, Olivier Le Maître, Ibrahim Hoteit, and Omar M. Knio.

Appendix A. VPP deterministic model

In this appendix, the deterministic model of the VPP and its data are presented. The goal of the problem is to maximize the operational profit defined as

$$P = \max \sum_{f} \sum_{j} \left[\left(\lambda_{f,j}^{sell} f_{f,j}^{sell} - \lambda_{f,j}^{buy} f_{f,j}^{buy} \right) D_f \right] + \sum_{t} \left[\lambda_t \left(p_t^{sell} - p_t^{buy} \right) \right] - cop, \quad \text{(Appendix A1)}$$

$$\sum_{i \in TH} p_{i,t} + \sum_{i \in HY} ptb_{i,t} + p_t^{buy} + \sum_{f} \sum_{j} f_{f,j}^{buy} + w_t = \sum_{i \in HY} pp_{i,t} + p_t^{sell} + \sum_{f} \sum_{j} f_{f,j}^{sell}, \quad \forall t, \quad \text{(Appendix A2)}$$

$$f_{f,j}^{buy} \leq F_{f,j}^{buy} \quad \forall f, j, \quad \text{(Appendix A3a)}$$

$$f_{f,j}^{sell} \leq F_{f,j}^{sell} \quad \forall f, j, \quad \text{(Appendix A3b)}$$

$$\sum_{j} f_{f,j}^{buy} \leq \sum_{j} F_{f,j}^{buy} y_f^{buy} \quad \forall f, \quad \text{(Appendix A3c)}$$

$$\sum_{j} f_{f,j}^{sell} \leq \sum_{j} F_{f,j}^{sell} y_f^{sell} \quad \forall f, \quad \text{(Appendix A3d)}$$

$$y_f^{sell} + y_f^{buy} \leq 1 \quad \forall f, \quad \text{(Appendix A3e)}$$

$$\sum_{tt \geq t - UT_i + 1, tt \leq t} u_{i,tt}^{up} \leq u_{i,t} \quad \forall i, t \geq LM_i + 1, \quad \text{(Appendix A4)}$$

$$u_{i,t} + \sum_{tt \geq t - DT_i + 1, tt \leq t} u_{i,tt}^{dn} \leq 1 \quad \forall i, t \geq FM_i + 1, \quad \text{(Appendix A5)}$$

$$P_i^l u_{i,t} \leq p_{i,t} \leq P_i^u u_{i,t} \quad \forall i \in TH, t, \quad \text{(Appendix A6)}$$

$$p_{i,t} \leq P0_i + RU_i U0_i + SU_i u_{i,t}^{up} \quad \forall i \in TH, t = 1, \quad \text{(Appendix A7)}$$

$$p_{i,t} - p_{i,t-1} \leq RU_i u_{i,t-1} + SU_i u_{i,t}^{up} \quad \forall i \in TH, t > 1, \quad \text{(Appendix A8)}$$

$$p_{i,t-1} - p_{i,t} \leq RD_i u_{i,t} + SD_i u_{i,t}^{dn} \quad \forall i \in TH, t > 1, \quad \text{(Appendix A9)}$$

$$cop = \sum_{i \in TH} \sum_{t} cu_{i,t} + A_i u_{i,t} + B_i p_{i,t} + cd_{i,t}, \quad \text{(Appendix A10)}$$

$$cd_{i,t} \geq DC_i \left(1 - u_{i,t}\right) \quad \forall i, t = 1, T_i^I > 0, \quad \text{(Appendix A11)}$$

$$cd_{i,t} \geq u_{i,t}^{dn} DC_i \quad \forall i, t \geq 2, \quad \text{(Appendix A12)}$$

$$cu_{i,t} \geq u_{i,t}^{up} HS_i, \quad \forall i, t, \quad \text{(Appendix A13)}$$

$$cu_{i,t} \geq \left(u_{i,t} - \sum_{tt \geq t - (DT_i + T_i^c + 1)}^{t-1} u_{i,tt} \right) CS_i \quad \forall i, t > TD_i + T_i^c, \quad \text{(Appendix A14)}$$

$$cu_{i,t} \geq \left(u_{i,t} - \sum_{tt < t} u_{i,tt} \right) CS_i, \ \forall i, T_i^I < 0, \ \left(TD_i + T_i^c + T_i^I + 1 \right) < t \leq \left(TD_i + T_i^c \right), \quad \text{(Appendix A15)}$$

$$1 - u_{i,t} + u_{i,t}^{up} - u_{i,t}^{dn} = 0, \quad \forall i, t = 1, T^I > 0, \quad \text{(Appendix A16)}$$

$$-u_{i,t} + u_{i,t}^{up} - u_{i,t}^{dn} = 0, \quad \forall i, t = 1, T^I < 0, \quad \text{(Appendix A17)}$$

$$u_{i,t-1} - u_{i,t} + u_{i,t}^{up} - u_{i,t}^{dn} = 0, \quad \forall i, t > 1, \quad \text{(Appendix A18)}$$

$$v_{i,t} - G(-q_{i,t} + qp_{i,t}) = V0_i + GQ_i^{in} \quad \forall i \in HY, t = 1, \quad \text{(Appendix A19)}$$

$$v_{i,t} - v_{i,t-1} - G(-q_{i,t} + qp_{i,t}) = GQ_i^{in} \quad \forall i \in HY, t > 1, \quad \text{(Appendix A20)}$$

$$ptb_{i,t} - K_i^{\intercal} q_{i,t} H_i = 0 \quad \forall i \in HY, t, \quad \text{(Appendix A21)}$$

$$pp_{i,t} - K_i^p qp_{i,t} H_i = 0 \quad \forall i \in HY, t, \quad \text{(Appendix A22)}$$

$$q_{i,t} \leq Q_i^u \quad \forall i \in HY, t, \quad \text{(Appendix A23)}$$

$$qp_{i,t} \leq Q_i^u \quad \forall i \in HY, t, \quad \text{(Appendix A24)}$$

$$V_{i,t}^l \leq v_{i,t} \leq V_{i,t}^u \quad \forall i \in HY, t, \quad \text{(Appendix A25)}$$

$$v_{i,t} \geq V_i^E \quad \forall i \in HY, t = tf, \tag{Appendix A26}$$
$$f_{f,j}^{buy}, f_{f,j}^{sell} \geq 0, \quad \forall f, j, \tag{Appendix A27}$$
$$y_f^{sell}, y_f^{buy} \in \{0,1\}, \quad \forall f, \tag{Appendix A28}$$
$$u_{i,t}, u_{i,t}^{up}, u_{i,t}^{dn} \in \{0,1\}, \quad \forall i \in TH, t, \tag{Appendix A29}$$
$$p_{i,t} \geq 0, \quad \forall i \in TH, t, \tag{Appendix A30}$$
$$p_t^{sell}, p_t^{buy} \geq 0, \quad \forall t, \tag{Appendix A31}$$
$$ptb_{i,t}, pp_{i,t}, q_{i,t}, qp_{i,t}, v_{i,t} \geq 0, \quad \forall i \in HY, t. \tag{Appendix A32}$$

The contracts in (Appendix A3a) to (Appendix A3e) are modeled based on Conejo et al. [62]. The thermal unit considers: a) minimum up and down-time limits, (Appendix A4) and (Appendix A5); b) bounds on the minimum and maximum power output, (Appendix A6) to (Appendix A9); c) fixed and variable costs of power generation, startups, and shutdowns, (Appendix A10) to (Appendix A15); and d) logical relations between the binary variables that define the states of the unit, (Appendix A16) to (Appendix A18). $LM_i = \min\{|T|, U_i\}$, U_i is the number of hours the thermal unit i needs to be On at the beginning of the time horizon, $FM_i = \min\{|T|, DMi\}$, DM_i is the number of hours the thermal unit needs to be Off at the beginning of the time horizon. The pumped-storage hydro plant model considers: a) water mass balances for the reservoir at the end of each period, (Appendix A19), and (Appendix A20); and b) power generation and consumption functions, (Appendix A21), and (Appendix A22). The parameters of the thermal unit, pump-storage hydro plant and contracts are shown in the Tables S1, S2, and S3.

Table S1: Base data for the thermal generators.

Case	P^L (MW)	P^U (MW)	UT (h)	DT (h)	T^c (h)	T^I (h)	SU/SD (MW/h)	RU/RD (MW/h)	A_i (€/h)	B_i (€/MWh)	HS_i (€/h)	CS_i (€/h)
1	150	455	8	8	5	8	150	91	1000	16.19	4500	9000
2	10	55	1	1	0	-1	10	11	660	25.92	30	60

Table S2: Data for the hydro plant.

Case	H_i (m)	K_i^p (MWs/m^3)	$K_i^\intercal$ (MWs/m^3)	Q_i^{in} (m^3/s)	Q_i^u (m^3/s)	$V0_i$ (Hm^3)	V_i^u (Hm^3)	V_i^l (Hm^3)
1 and 2	113	0.99	1	0.1	46.5	84	560	10

Table S3: Price of the energy (€/MWh) for two contracts for each week.

			Sell blocks			Buy blocks		
Week	Contract	Block Size	3	2	1	1	2	3
1	1	50	64.77	59.77	54.77	54.77	49.77	44.77
1	2	55	58.29	53.79	49.29	49.29	44.79	40.29

Nomenclature

Sets

F	Forward contracts
HY	Hydro pump-storage generation units
J	Blocks of the forward contracts
I	Generating units
T	Time periods
TH	Thermal generation units

Parameters

A_i, B_i	Production cost function coefficients for unit i (€/h), (€/MWh)
CS_i	Cold start-up cost of unit i (€/h)
DM_i	Number of periods unit i must be off at the beginning of the time horizon
D_f	Time periods spanned by contract f
DC_i	Shut-down cost (€)
DT_i	Minimum down-time of unit i (h)
FM_i	Minimum number of periods a unit i must be off at the beginning of the time horizon
HS_i	Hot start cost of unit i (€/h)
LM_i	Minimum number of periods a unit i must be on at the beginning of the time horizon
P_i^l	Minimum power output of unit i (MW)
P_i^u	Maximum power output of unit i (MW)
$P0_i$	Power produced at t=0 by unit i (MW)
RD_i	Maximum ramp-down rate of unit i (MW)
RU_i	Maximum ramp-up rate of unit i (MW)
SD_i	Maximum shutdown rate of unit i (MW)
SR_t	Spinning reserve for period t (MW)
SU_i	Maximum start-up rate of unit i (MW)
U_i	Number of periods unit i must be on at the beginning of the time horizon
$U0_i$	Initial state of unit i {on,off}={1,0}
UT_i	Minimum up-time of unit i (h)
T_i^c	Cold start hours of unit i (h)
T_i^I	Initial status of unit i (h)
G	Conversion factor between Hm^3 and m^3/s in one hour
H_i	Water head in plant i (m)
K_i^p	Power consumption factor
$K_i^{\intercal}$	Power generation factor
Q_i^{in}	Natural inflow of water for plant i (m^3/s)
Q_i^u	Maximum turbined and pumped flow of water for plant i (m^3/s)
V_i^u	Maximum volume of water in the reservoir of plant i (Hm^3)
V_i^l	Minimum volume of water in the reservoir of plant i (Hm^3)
V_i^E	Minimum volume of water in the reservoir of plant i at the end of the horizon (Hm^3)
$\lambda_{f,j}^{buy}$	Energy price of buying block j of forward contract f (€/MWh)

$\lambda_{f,j}^{sell}$ Energy price of selling block j of forward contract f (€/MWh)

Continuous variables

$cd_{i,t}$ Shut-down cost of unit i in period t (€)

cop Total startup, shutdown, production, and online cost of unit i (€)

cp Total startup, shutdown and online cost of unit i (€)

$cu_{i,t}$ Startup cost of unit i in period t (€)

$f_{f,j}^{buy}$ Power bought through block j of forward contract f (MW)

$f_{f,j}^{sell}$ Power sold through block j of forward contract f (MW)

P Operational profit of the producer per week (€)

$p_{i,t}$ Power output of unit i in period t (MW)

$p_{i,t}^{buy}$ Power bought in the pool in period t (MW)

p_t^{sell} Power sold in the pool in period t (MW)

$ptb_{i,t}$ Power output of the pump-storage hydro plant i in period t (MW)

$pp_{i,t}$ Power consumption of the pumped-storage hydro plant i in period t (MW)

$q_{i,t}$ Turbined flow of water in plant i in period t (m^3/s)

$qp_{i,t}$ Pumped flow of water in plant i in period t (m^3/s)

$v_{i,t}$ Volume of water stored in the reservoir of plant i (Hm^3)

Binary variables

$u_{i,t}$ On/off status of unit i in period t

$u_{i,t}^{up}$ Startup status of unit i in period t

$u_{i,t}^{dn}$ Shutdown status of unit i in period t

y_f^{buy} Selection of forward contract f to buy energy

y_f^{sell} Selection of forward contract f to sell energy

Random variables

w_t Wind power output in period t (MW)

λ_t Pool price in period t (€/MWh)

Appendix B. Uncertainty characterization for wind power and electricity prices.

The samples of the electricity price forecast are generated using an Auto-Regressive Integrated Moving Average model [63]. Considering the specific weeks of interest, we use the previous 12 weeks of historical time series of the electricity prices from the Iberian Electricity Market [64] to fit the ARIMA model. Then, we generate samples of the electricity forecast by sampling the error term of the ARIMA model for the 168 hours of interest. The structure of the ARIMA model is described in detail in Lima et al. [6], and additional information about these models applied to electricity prices can be found in Conejo et al. [65] and Weron [66].

To generate the wind power samples, we start with a wind speed ensemble forecast, with 51 members, for 168 hours for the location of the wind farm, obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF); see Lima et al. [6] for further details on the wind ensemble generation. The wind speed is converted to wind power using a wind-power curve for a turbine with a rated power of 2.35 MW. Based on the wind ensemble and a truncated Karhunen-Loève Expansion (KLE), we generate the samples of the wind power; see Le Maître and Knio [67, Section 2.1] for a presentation of the KLE and Lima et al. [55] for the description of the application of the KLE to the wind ensemble.

Concerning the ARO, the uncertainty sets are built using the same underlying data that are used to generate the samples for the SP. The uncertainty set for the electricity prices is also based on the ARIMA model, namely the point forecast of the ARIMA model is the forecast value, $\tilde{c}^f$ in (8), and the upper and lower bounds of the uncertainty set are based on the 95% forecast prediction error of the ARIMA model. In Figure 2, we provide the point forecast and bounds of the uncertainty set of the electricity prices for the two weeks, and in Figure 3, we show two samples: one with 10 elements and another with 100 elements.

The uncertainty set for the wind power is defined as the envelope that encompasses all the forecasts in the wind power ensemble that is used to build the KLE. In Figure 4, we present the original wind power ensemble and the uncertainty set elements, and in Figure 5, we show two samples: one with 10 elements and another with 100 elements. An important remark is that the uncertainty sets cut-off extreme points of the wind power and electricity price forecasts. These uncertainty sets constructions based on the ARIMA forecast errors and wind envelopes are more elaborated than the ones proposed in Lima et al. [19].

For Week 1, the electricity price and wind power forecast variability increase with time, which it is reflected in the range of the uncertainty sets, see Figures 2a and 4a. For Week 2, the electricity price forecast variability does not increase with time, see Figures 2b, while the wind power forecast variability increases significantly after 48 hours, see Figure 4c.

In the sampling generation and construction of the uncertainty sets, we do not consider a correlation between the electricity prices and the wind speed explicitly at the specific wind farm location because of the small capacity of the VPP compared with the power system where it is integrated. The study of this specificity is not the main purpose of this work. However, the models and methods proposed here can be valuable tools to analyze the impact of assumptions on potential correlations.

Appendix C. Stochastic model

In this appendix, we present a general stochastic formulation that captures the main characteristics of the VPP model presented in Appendix Appendix A:

$$\begin{aligned}
\max_{x^+,x^-,v,z,y^+(\theta),y^-(\theta),s(\theta),r(\theta)} \quad & \psi\left[f(x^+,x^-,z,y^+(\theta),y^-(\theta),s(\theta),\theta)\right] \\
\text{s.t.} \quad & A^+x^+ + A^-x^- + Bv \leq b \\
& Ez \leq g \\
& Cz + Ds(\theta) \leq d, \quad \theta \in \Theta \\
& s(\theta) - y^+(\theta) + y^-(\theta) + r(\theta) - e^\top x^+ + e^\top x^- + h(\theta) = 0, \quad \theta \in \Theta \\
& Jr(\theta) \leq j, \quad \theta \in \Theta \\
& x^+, x^- \in \mathbb{R}^{n_1}_+, v \in \mathbb{B}^{n_1}, z \in \mathbb{B}^{n_2}, y^+(\theta), y^-(\theta), s(\theta) \in \mathbb{R}^{n_2}_+, r(\theta) \in \mathbb{R}^{n_2},
\end{aligned} \tag{Appendix C1}$$

with the function f representing the random profit defined as

$$f(x^+,x^-,z,y^+(\theta),y^-(\theta),s(\theta),\theta) := (c^+)^\top x^+ + (c^-)^\top x^- + \bar{c}^\top z + \tilde{c}^\top(\theta)y^+(\theta) - \tilde{c}^\top(\theta)y^-(\theta) + \underline{c}^\top s(\theta), \tag{Appendix C2}$$

where $\mathbb{B} := \{0,1\}$, v represents the binary variables associated with the blocks in the contracts, $h(\theta) \in \mathbb{R}^{n_1}$ symbolizes the uncertain wind power, $A^+, A^- \in \mathbb{R}^{m_1 \times n_1}$, $B \in \mathbb{R}^{m_1 \times n_1}$, $C \in \mathbb{R}^{m_2 \times n_2}$, $D \in \mathbb{R}^{m_2 \times n_2}$, $E \in \mathbb{R}^{m_3 \times n_2}$, $J \in \mathbb{R}^{m_4 \times n_2}$ are matrices with known parameters, and $b \in \mathbb{R}^{m_1}$, $d \in \mathbb{R}^{m_2}$, $e \in \mathbb{R}^{n_1}$, $g \in \mathbb{R}^{m_3}$, $j \in \mathbb{R}^{m_4}$ are vectors with known parameters. In (Appendix C1), x^+, x^-, v, z repesent the first-stage variables denoting the power to sell by contract, the power to buy by contract, the binary variables associated with the contracts, and the binary variables associated with the thermal unit commitment, respectively. These variables correspond to $f^{buy}_{f,j}$, $f^{sell}_{f,j}$, $u_{i,t}$, $u^{up}_{i,t}$, $u^{dn}_{i,t}$, y^{buy}_f, y^{sell}_f used in the deterministic model. The second-stage variables are $y^+(\theta)$, $y^-(\theta)$, $s(\theta)$, and $r(\theta)$, which stand for the power to sell in the electricity market pool, the power to buy in the electricity market pool, the variables related with power produced by the thermal unit, and the variables related with the pump-storage hydro plant. These variables correspond to $p^{buy}_{i,t}$, p^{sell}_t, $p_{i,t}$, $ptb_{i,t}$, $pp_{i,t}$, $q_{i,t}$, $qp_{i,t}$, $v_{i,t}$ used in the deterministic model. The first constraint in (Appendix C1) models the forward contracts setup, the second the logic constraints defined by the binary variables related with the thermal unit commitment, the third establishes the relationship between the binary variables and continuous variables used to model the commitment of the thermal unit, the fourth the energy balance, and the fifth captures the relations between the variables that model the operation of the pump-storage hydro plant. The formulation described here is the basis for the compact formulation introduced in Section 3.

Appendix D. Comparison between the dual and primal variants of the ARO decomposition algorithm

In Case 1 (Week 1 and Week 2), both variants of the ARO decomposition algorithm have a similar performance; see the Tables S4 and S6. However, in Case 2, Week 1, for $\Gamma = 10$ there is a clear difference between the computational time required by the dual and the primal variants: 1039 s vs. 235 s; see Table S5. For the remaining values of Γ, the gap between the bounds is reduced faster with the primal variant. We show and example of these faster reductions in Figures S1 and S2. Also, in Case 2, Week 2, the performance of the primal is also superior for $\Gamma = 10$. The primal version requires 30 s to achieve a gap of 0.00, whereas the dual terminates with a gap of 11.75% after 10,929 s. In addition, the primal version achieves a gap of 0.00 for $\Gamma = 150$ and $\Gamma = 168$ in 1820 s and 1,079 s, respectively. For the same valus of Γ, the dual variant terminates with a gap of 77.47% and 33.81%, in more than 10,800 s. In Figures S3 and S4, the faster evolution of the bounds of the primal variant is contrasted with the stagnant bounds of the dual variant, for Case 2, Week 2, and $\Gamma = 168$.

Table S4: Case 1, Week 1. Performance of the ARO decomposition algorithm with three versions of the master problem: D - optimality cut with dual variables; D+P - optimality cut with dual variables plus primal constraints and variables; P - primal constraints and variables.

	LB			ITER			GAP (%)			Time (s)		
Γ	D	P	D+P	D	P	D+P	D	P	D+P	D	P	D+P
0	2,937,740	2,937,740	2,937,740	2	2	2	0.00	0.00	0.00	1	1	1
10	2,789,821	2,789,821	2,789,821	3	3	3	0.00	0.00	0.00	6	6	6
50	2,372,466	2,372,466	2,372,466	4	4	4	0.00	0.00	0.00	25	26	26
100	2,103,834	2,103,834	2,103,834	4	4	4	0.00	0.00	0.00	46	46	47
150	1,953,777	1,953,777	1,953,777	4	4	4	0.00	0.00	0.00	13	16	12
168	1,912,678	1,912,678	1,912,678	3	3	3	0.00	0.00	0.00	2	2	3

Table S5: Case 2, Week 1. Performance of the ARO decomposition algorithm with three versions of the master problem: D - optimality cut with dual variables; D+P - optimality cut with dual variables plus primal constraints and variables; P - primal constraints and variables.

	LB			ITER			GAP (%)			Time (s)		
Γ	D	P	D+P	D	P	D+P	D	P	D+P	D	P	D+P
0	309,568	309,568	309,568	7	2	2	0.00	0.00	0.00	3	1	3
10	277,168	277,168	277,168	22	5	5	0.00	0.00	0.00	1,039	235	235
50	220,130	220,225	220,316	106	36	30	1.24	1.21	1.16	11,147	10,822	10,938
100	210,219	211,019	210,480	70	29	30	1.41	1.02	1.29	11,111	10,884	10,939
150	208,918	209,426	208,815	38	29	29	1.29	1.05	1.34	10,825	10,892	10,875
168	208,787	209,377	208,995	39	28	28	1.33	1.04	1.23	10,881	11,269	11,267

Table S6: Case 1, Week 2. Performance of the ARO decomposition algorithm with three versions of the master problem: D - optimality cut with dual variables; D+P - optimality cut with dual variables plus primal constraints and variables; P - primal constraints and variables.

	LB			ITER			GAP (%)			Time (s)		
Γ	D	P	D+P	D	P	D+P	D	P	D+P	D	P	D+P
0	2,381,565	2,381,565	2,381,565	2	2	2	0.00	0.00	0.00	2	1	3
10	2,207,579	2,207,579	2,207,579	4	4	4	0.00	0.00	0.00	4	5	5
50	1,745,062	1,745,062	1,745,062	3	3	3	0.00	0.00	0.00	9	10	10
100	1,452,951	1,452,951	1,452,951	4	4	4	0.00	0.00	0.00	42	43	44
150	1,332,694	1,332,694	1,332,694	4	4	4	0.00	0.00	0.00	6	3	3
168	1,318,392	1,318,392	1,318,392	6	5	5	0.00	0.00	0.00	3	4	4

Table S7: Case 2, Week 2. Performance of the ARO decomposition algorithm with three versions of the master problem: D - optimality cut with dual variables; D+P - optimality cut with dual variables plus primal constraints and variables; P - primal constraints and variables.

	LB			ITER			GAP (%)			Time (s)		
Γ	D	P	D+P	D	P	D+P	D	P	D+P	D	P	D+P
0	333,371	340,353	340,353	266	2	2	4.84	0.00	0.00	10,832	1	5
10	276,609	285,064	285,064	104	4	4	11.75	0.00	0.00	10,929	30	30
50	111,136	184,801	184,875	19	30	30	111.15	2.46	2.41	12,429	11,288	11,257
100	18,621	155,500	155,470	36	31	31	1,003.44	0.15	0.17	12,017	11,138	11,110
150	108,257	151,925	151,925	147	17	13	77.47	0.00	0.00	10,832	1,820	1,889
168	136,363	151,821	151,821	242	11	13	33.81	0.00	0.00	10,851	1,079	1,179

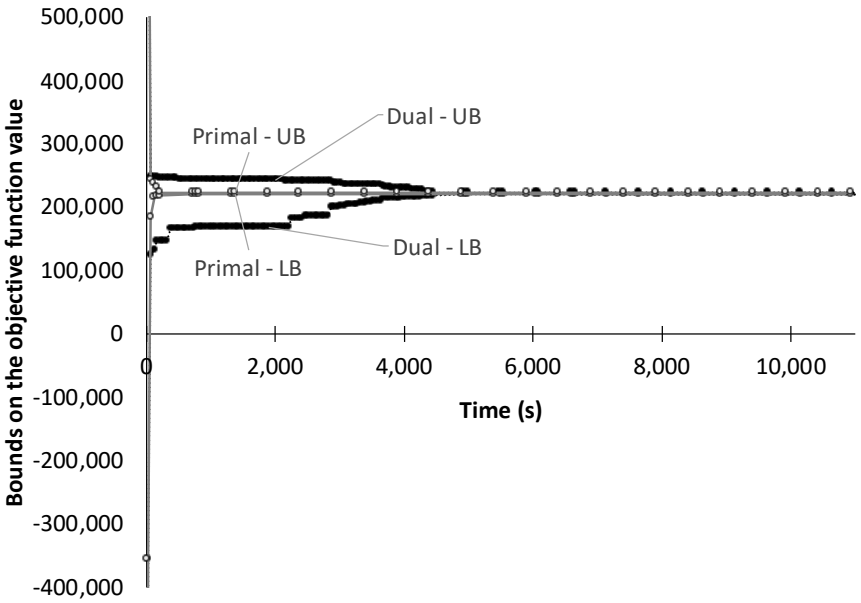


Figure S1: Case 2, Week 1. Profile of the upper and lower bounds on the optimal value of the original problem obtained with the dual and primal variants of the master problem within the ARO decomposition algorithm. $\Gamma = 50$.

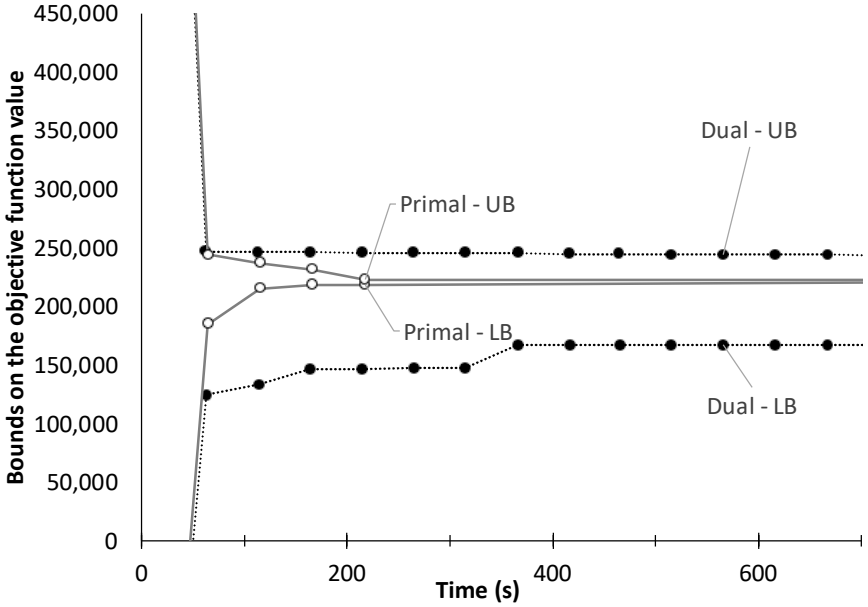


Figure S2: Case 2, Week 1. Zoom in of Figure S1.

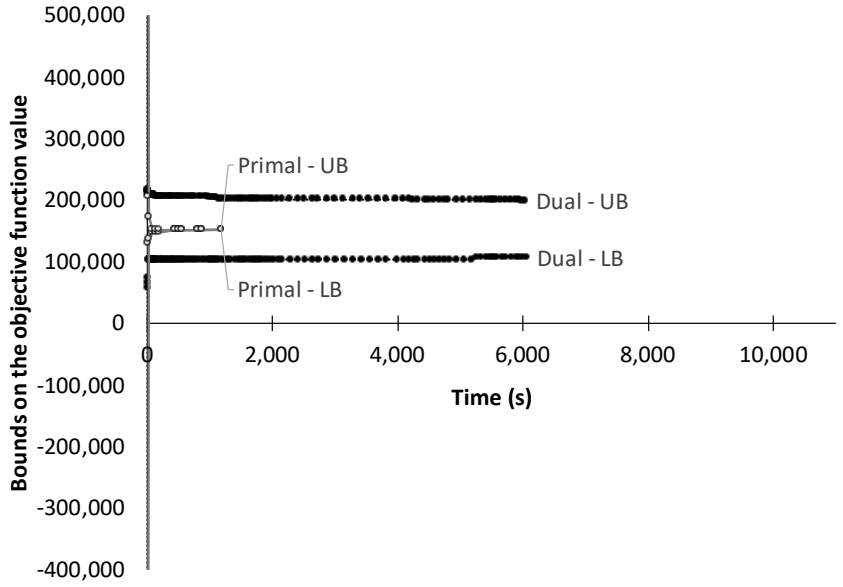


Figure S3: Case 2, Week 2. Profile of the upper and lower bounds on the optimal value of the original problem obtained with the dual and primal variants of the master problem within the ARO decomposition algorithm. $\Gamma = 168$.

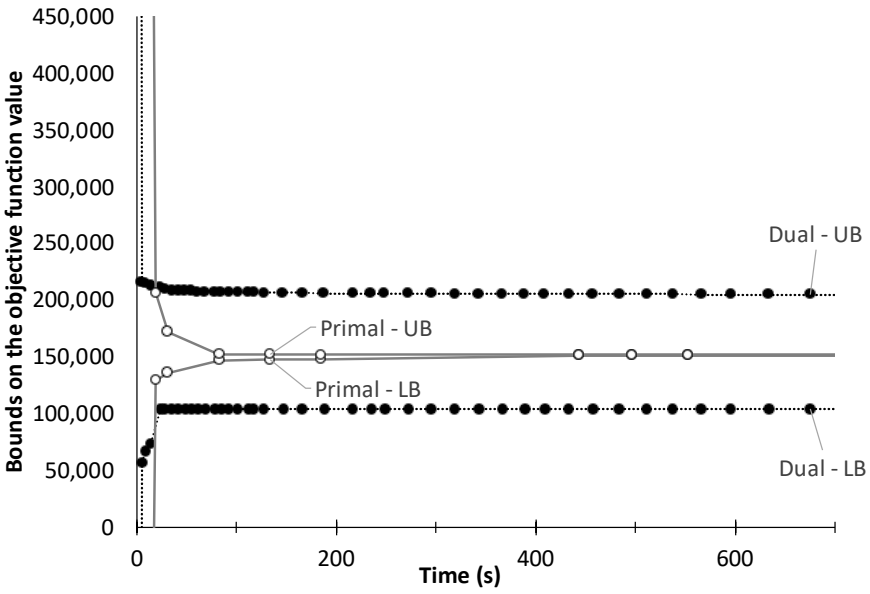


Figure S4: Case 2, Week 2. Zoom in of Figure S3.

Appendix E. Detailed results and discussion on the adaptive robust optimization and stochastic programming performance

Appendix E.1. Adaptive robust optimization performance.

We analyze the performance of the ARO decomposition algorithm based on each case and week, and the budget of uncertainty parameter, Γ. The computational times reported in the following subsections refer to one optimization run without including the computational time required by the bound estimation stage. We start by dividing the results by week, and then for each week we analyze each case. For the sake of compactness, we use ARO as an alias for ARO decomposition algorithm.

Appendix E.1.1. Week 1.

Figure S5a shows that Case 1 leads to a better performance of the ARO than Case 2. Also, the results of Case 1 show that the performance depends on Γ.

More specifically, in Case 1, the ARO is relatively fast with $\Gamma = 168$, requiring 3 s to meet the gap stopping criterion, whereas with $\Gamma = 100$, it requires 47 s.

In Case 2, the ARO can only meet the gap stopping criterion for $\Gamma = 10$. However, it requires 235 s, against the 6 s in Case 1. Consequently, the final number of iterations is higher in Case 2 than in Case 1; see Tables S8 and S9.

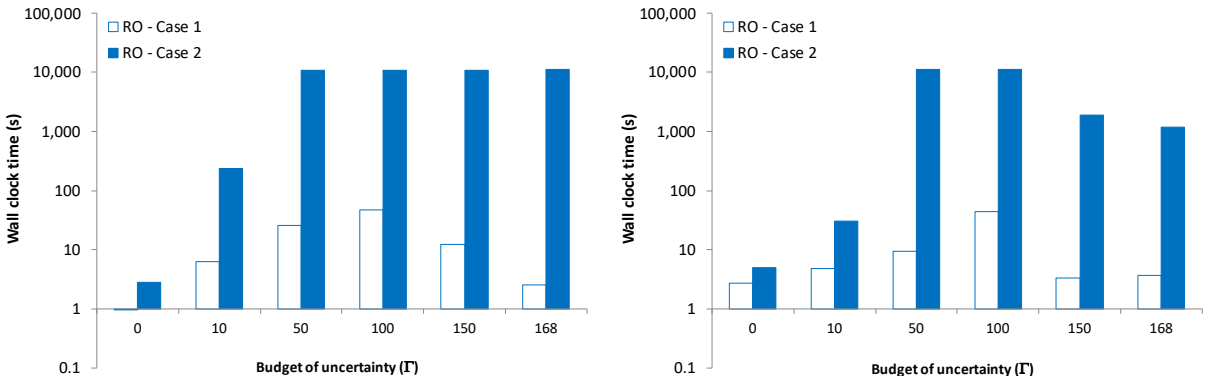
Appendix E.1.2. Week 2.

In Case 1, the performance of ARO has a similar trend compared to Week 1; see Figures S5a and S5b, and Tables S8 and S10. Also, the performance of the ARO with Case 1 is better than with Case 2. Figure S5b and Table S11 show that in Case 2 with $\Gamma = \{150, 168\}$, the gap stopping criterion is met on 1,889 s and 1,179 s, which does not occur for these Γ in Week 1. But still, these computational times are three orders of magnitude higher than in Case 1; see Figure S5b.

Appendix E.1.3. Discussion on the ARO performance.

Overall, the technical characteristics of the thermal unit induce a clear-cut difference in the performance of the ARO for the same uncertainty data. This is supported by the differences between the results of Case 1 and Case 2, for the same week. The value of Γ also influences the results, with the extremes of Γ leading to more manageable problems.

A detailed analysis of the performance of the ARO model reveals that the MILP solver does not close the optimality gap for some MILP subproblems. Therefore, based on the adaptive maximum time strategy, the maximum time for the MILP solver is increased from 50 s to 500 s because the upper bound does not



(a) Week 1. (b) Week 2

Figure S5: Wall clock time for the ARO as a function of the budget of uncertainty parameter.

Table S8: Case 1. Week 1. Performance of the decomposition algorithms averaged over $M = 30$ optimization replicas.

	Risk parameters	N	SR	ITER	GAP (%)	Time (s)
SP	$\beta = 0$	10	30	2.0	0.0000	1
		50	30	2.0	0.0000	2
		100	30	2.0	0.0000	3
		500	30	2.0	0.0000	4
		5000	30	2.0	0.0000	43
SP	$\beta = 1$, $\alpha = 0.95$	50	30	2.0	0.0000	2
		100	30	2.0	0.0000	3
		500	30	2.0	0.0000	5
		5000	30	2.0	0.0000	49
SP	$\beta = 1$, $\alpha = 1$	10	30	2.1	0.0000	1
		50	30	2.3	0.0000	2
		100	30	2.7	0.0000	5
		500	30	5.7	0.0000	31
		5000	26	9.3	0.0530	2,556
RO	Γ					
	0	-	1	2	0.0000	1
	10	-	1	3	0.0000	6
	50	-	1	4	0.0000	26
	100	-	1	4	0.0000	47
	150	-	1	4	0.0000	12
	168	-	1	3	0.0000	3

N - sample size used in each optimization replica; SR - number of successful optimization replications that meet the gap stop criterion (out of 30); ITER - average number of iterations of the decomposition algorithm; GAP - average gap between the lower bound and upper bound of the decomposition algorithms; Time - average wall clock time.

Table S9: Case 2. Week 1. Performance of the decomposition algorithms averaged over $M = 30$ optimization replicas.

	Risk parameters	N	SR	ITER	GAP (%)	Time (s)
SP	$\beta = 0$	10	30	6.3	0.0000	3
		50	30	6.0	0.0000	5
		100	30	6.0	0.0000	9
		500	30	6.0	0.0000	17
		5000	30	6.0	0.0000	180
SP	$\beta = 1$, $\alpha = 0.95$	50	30	20.5	0.0000	60
		100	30	9.0	0.0000	22
		500	30	6.0	0.0000	26
		5000	30	6.0	0.0000	294
SP	$\beta = 1$, $\alpha = 1$	10	29	103.2	0.1279	1,399
		50	30	39.8	0.0000	206
		100	30	30.4	0.0000	230
		500	30	18.6	0.0000	461
		5000	28	11.9	0.0263	2,006
RO	Γ					
	0	-	1	2	0.0000	3
	10	-	1	5	0.0000	235
	50	-	0	30	1.1588	10,938
	100	-	0	30	1.2893	10,939
	150	-	0	29	1.3357	10,875
	168	-	0	28	1.2290	11,267

Table S10: Case 1. Week 2. Performance of the decomposition algorithms averaged over $M = 30$ optimization replicas.

	Risk parameters	N	SR	ITER	GAP (%)	Time (s)
SP	$\beta = 0$	10	30	2.0	0.0000	1
		50	30	2.0	0.0000	2
		100	30	2.0	0.0000	3
		500	30	2.0	0.0000	5
		5000	30	2.0	0.0000	44
SP	$\beta = 1$, $\alpha = 0.95$	50	30	2.3	0.0000	2
		100	30	2.0	0.0000	3
		500	30	2.0	0.0000	5
		5000	30	2.0	0.0000	111
SP	$\beta = 1$, $\alpha = 1$	10	30	3.8	0.0000	2
		50	30	3.3	0.0000	3
		100	30	5.8	0.0000	13
		500	30	4.4	0.0000	22
		5000	30	5.5	0.0000	906
RO	Γ					
	0	-	1	2	0.0000	3
	10	-	1	4	0.0000	5
	50	-	1	3	0.0000	10
	100	-	1	4	0.0000	44
	150	-	1	4	0.0000	3
	168	-	1	5	0.0000	4

N - sample size used in each optimization replica; SR - number of successful optimization replications that meet the gap stop criterion (out of 30); ITER - average number of iterations of the decomposition algorithm; GAP - average gap between the lower bound and upper bound of the decomposition algorithms; Time - average wall clock time.

Table S11: Case 2. Week 2. Performance of the decomposition algorithms averaged over $M = 30$ optimization replicas.

	Risk parameters	N	SR	ITER	GAP (%)	Time (s)
SP	$\beta = 0$	10	0	373.9	2.8711	10,849
		50	0	335.6	2.4068	10,838
		100	0	308.1	2.6011	10,851
		500	0	212.3	3.3464	10,849
		5000	0	65.8	5.6553	10,933
SP	$\beta = 1$, $\alpha = 0.95$	50	0	30.6	29.9756	11,354
		100	0	28.2	32.9694	11,322
		500	0	23.1	35.9119	11,408
		5000	0	10.5	20.5594	15,658
SP	$\beta = 1$, $\alpha = 1$	10	6	53.6	22.3870	9,741
		50	3	27.9	29.0440	10,376
		100	0	27.1	35.4289	11,364
		500	0	20.4	40.2817	11,587
		5000	0	8.5	37.8558	15,415
RO	Γ					
	0	-	1	2	0.0000	5
	10	-	1	4	0.0000	30
	50	-	0	30	2.4146	11,257
	100	-	0	31	0.1739	11,110
	150	-	1	13	0.0000	1,889
	168	-	1	13	0.0000	1,179

move between iterations. However, even with 500 s, the MILP solver does not close the gap of the MILP subproblems.

These results motivated an additional set of experiments with the ARO, where the maximum time for the MILP solver is increased to 5000 s, and the maximum time for the ARO is extended to 18,000 s. The solution approach described in Section 4.3 leads to a faster gap reduction, whereas the one with extended times leads to better final gaps in some cases. For additional results, we refer the reader to Appendix Appendix F.

Appendix E.2. Stochastic programming performance

We analyze the performance of the SP decomposition algorithm based on $M = 30$ optimization replications and over different sample sizes and risk parameters. The computational times reported in the following subsections do not include the computational time required by the bound estimation stage. For the sake of compactness, we use SP as an alias for SP decomposition algorithm.

Appendix E.2.1. Week 1.

In Case 1, computational times increase with sample size; see Figure S6a. However, the increment between $N = 10$ and $N = 500$ is only three seconds, for $\beta = 0$, and between $N = 50$ and $N = 500$, it is also three seconds, for $\beta = 1$, $\alpha = 0.95$. From another perspective, analyzing the computational times for $N = 5000$, we observe a steep increase from $\beta = 0$ to $\beta = 1$, $\alpha = 1$; see Figure S6a.

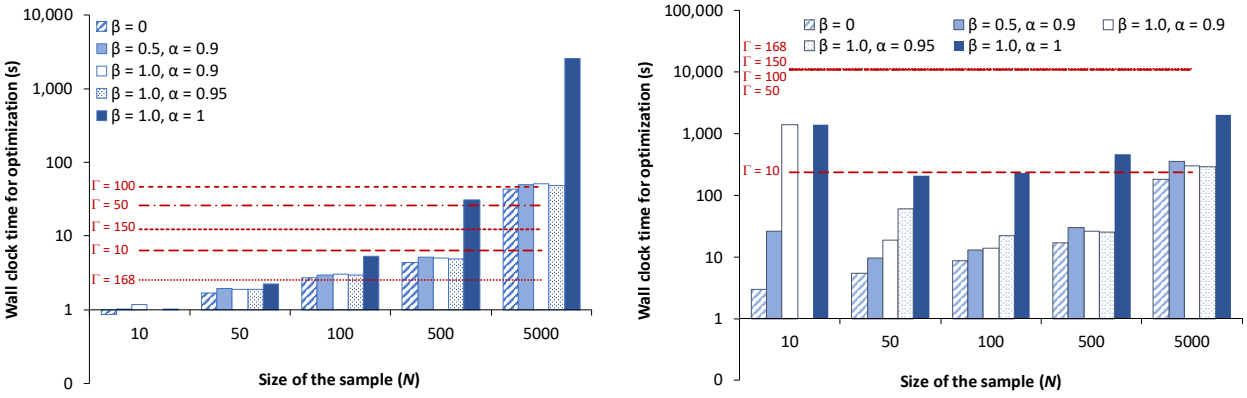
In Case 2, computational times do not increase monotonically with sample size, as in Case 1; see Figure S6b. For example for $\beta = 0.5$, $\alpha = 0.9$, the average computational time for $N = 10$ is greater than for $N = 50$ and $N = 100$. Also, the computational time for $\beta = 1$, $\alpha = 1$ is independent of the sample size. Overall, the SP model requires higher computational times in Case 2 than in Case 1.

Appendix E.2.2. Week 2.

In Case 1, the performance of the SP shows a similar trend to Week 1; see Figure S7a. However, in Case 2, the SP reaches the maximum time in most optimization replications without meeting the gap stopping criterion; see Figure S7b. There are only nine successful optimization replications out of 720 optimization runs (30 optimization replications times five sample sizes times five sets of risk parameters). The final gaps and iterations are presented in Table S11.

Appendix E.2.3. Discussion on the SP performance.

As discussed in the ARO section, the performance of the SP model is also sensitive to the technical characteristics of the thermal unit. Overall, the performance in Case 2 is worse than that in Case 1, within each week considered.



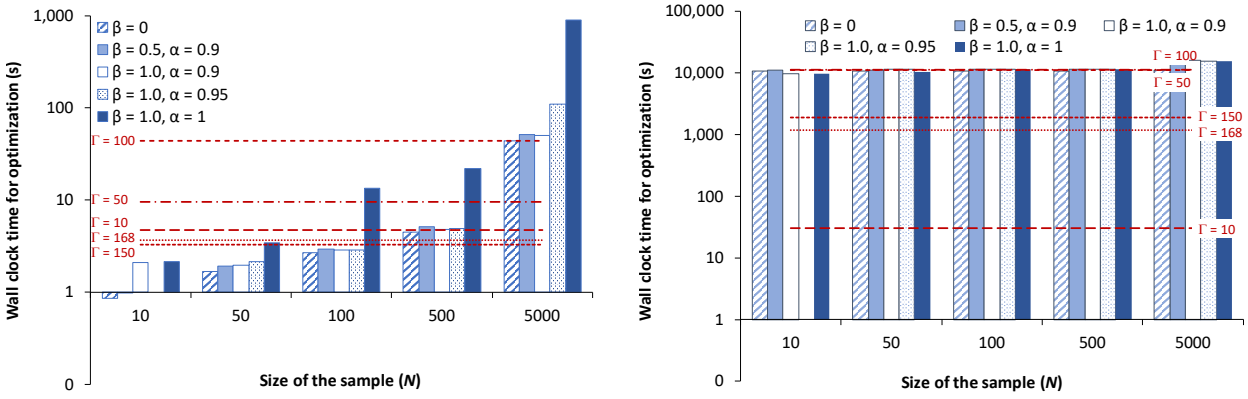
(a) Case 1.

(b) Case 2.

Figure S6: Week 1. Average wall clock time of $M = 30$ optimization replications of SP, and ARO.

The main difference between Case 1 and Case 2 is generation cost. In Case 2, generation costs are closer to electricity prices than in Case 1. Also, they are closer in Week 2 than in Week 1. These differences force the decomposition algorithm to perform more iterations in Case 2 than in Case 1 to find an optimal solution regarding the contract values and commitment of the thermal unit.

The results also indicate that risk-averse parameters lead to more involved problems. In Case 1, we observe that as the sample size increases, the problems with $\beta = 1$, $\alpha = 1$ require increasing time to solve. In Case 2, Week 1, the performance of the SP model for the extreme risk-averse parameters, $\beta = 1$, $\alpha = 1$, is independent of the sample size. In Week 2, the SP exhibits difficulties for all risk parameters and sample sizes.



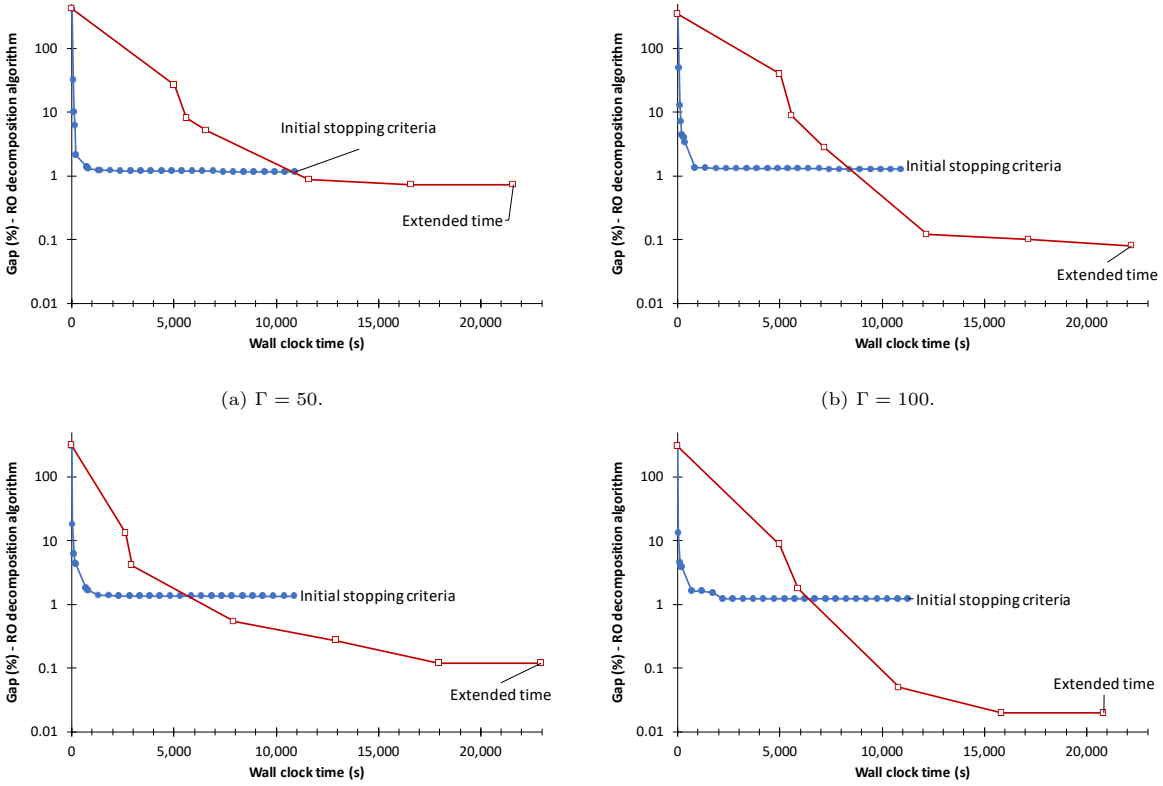
(a) Case 1.

(b) Case 2.

Figure S7: Week 2. Average wall clock time of $M = 30$ optimization replications of SP, and ARO.

Appendix F. Analysis of the relationship between the maximum time to solve the subproblem and the performance of the ARO decomposition algorithm

The performance of the ARO decomposition algorithm for Case 2, described in Section Appendix E.1, motivated a new set of experiments with the ARO decomposition algorithm, where the maximum time for the MILP solver to solve the MILP subproblem is increased to 5000 s and the maximum time for the ARO decomposition algorithm is extended to 18,000 s. For this set of experiments, the profile of the gap between the upper and lower bounds of the ARO decomposition algorithm is shown in Figures S8 and S9, for Week 1 and 2, respectively. These profiles are contrasted with the profiles from the original stopping criteria.



(a) $\Gamma = 50$.

(b) $\Gamma = 100$.

(c) $\Gamma = 150$.

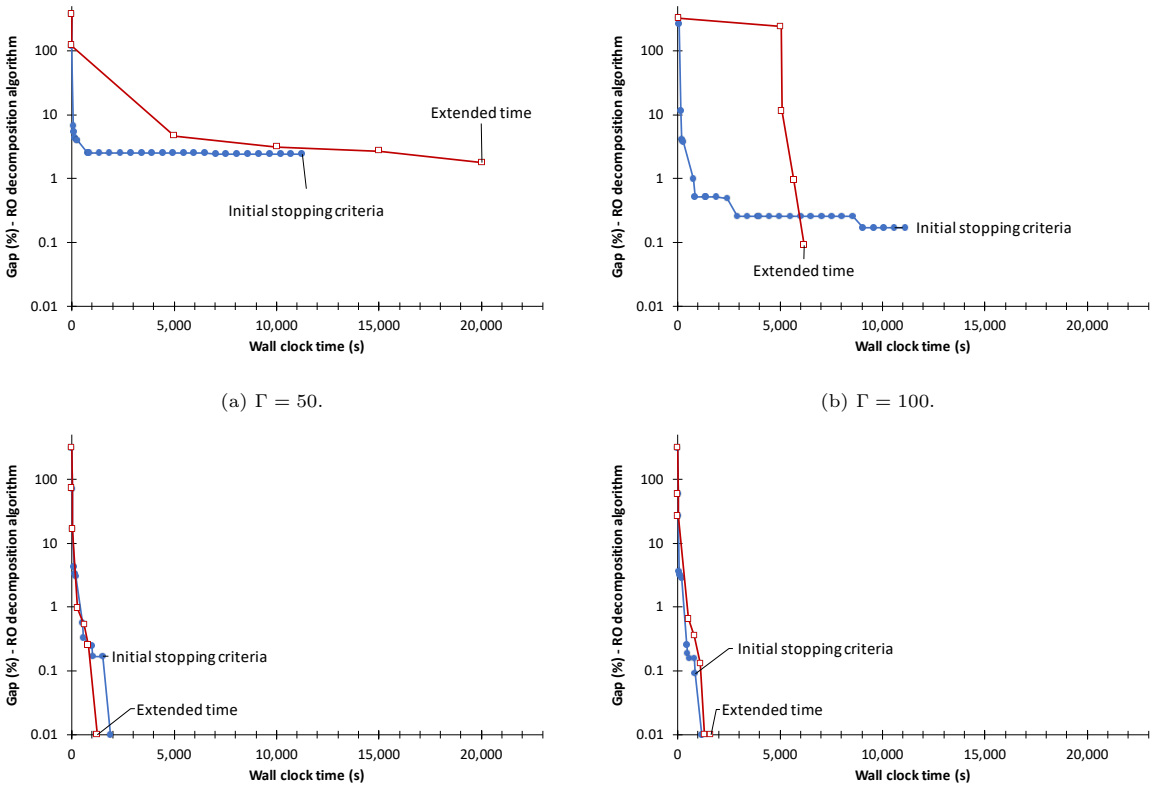
(d) $\Gamma = 168$.

Figure S8: Week 1. Profile of the gap in the ARO decomposition algorithm with the initial stopping criteria and with an extended time for the solution of the MILP subproblem and for the ARO decomposition algorithm.

These results provide the following insights: 1) the original stopping criteria lead to a faster reduction of the gap, whereas the new experiments lead to better final gaps; 2) some MILP subproblems are difficult to solve, by comparison with Case 1, requiring computational times above 5000 s to be solved to optimality; and 3) smaller gaps from the solution of the MILP subproblem lead to better bounds in the ARO decomposition algorithm.

We tested a number of strategies to improve the convergence of the ARO decomposition algorithm, namely, solving the relaxation of the master problem for some initial iterations, use different initial starting points, and a trust region method. However, a significant improvement was not achieved, which suggests that the master problem is not the limiting step, since these strategies are mainly related to this problem.

The results discussed using the two alternatives for the stopping criteria open the doors for research on more involved strategies to balance the time to solve the MILP subproblem.



(a) $\Gamma = 50$.

(b) $\Gamma = 100$.

(c) $\Gamma = 150$.

(d) $\Gamma = 168$.

Figure S9: Week 2. Profile of the gap in the ARO decomposition algorithm with the initial stopping criteria and with an extended time for the solution of the MILP subproblem and for the ARO decomposition algorithm.

Appendix G. Size of the master and subproblems

Table S12: Size of the master and subproblems for the first iteration and the increment per iteration.

		β	α	N	CONST	VAR	0-1 VAR	ΔCONST/I	ΔVAR/I
SP	Extensive	0	-	10	21,346	14,298	508	-	-
	Extensive	0	-	5,000	10,081,186	6,720,858	508	-	-
	Extensive	1	1	5,000	10,086,186	6,725,859	508	-	-
	Master	0	-	10	1,186	858	508	1	1
	problems	0	-	5000	1,186	858	508	1	1
		0.5	0.90	10	1,186	858	508	12	13
		1	0.90	10	1,186	858	508	11	12
		1	0.95	5000	1,186	858	508	5,001	5,002
		1	1	5000	1,186	858	508	5,001	5,002
	Subproblems†				1,345	2,353	0	0	0
RO	Master	-	-	-	1,186	858	508	2,017	1,513
	Subproblem†	-	-	-	2,355	3,025	672	0	0

† - The size presented is from the dual of the subproblem. N - sample size; CONST - number of constraints; VAR - number of total variables; 0-1 VAR - number of binary variables; ΔCONST/I - increment of number of constraints per iteration; ΔVAR/I - increment of number of variables per iteration.

Appendix H. Extended versions of Tables S8, S10, S9, and S11.

For the sake of completeness, this appendix contains results obtained with the SP approach with the risk parameters $\beta = 0.5$, $\alpha = 0.9$ and $\beta = 1$, $\alpha = 0.9$, which are not presented in Tables S8, S10, S9, and S11.

Table S13: Case 1. Week 1. Performance of the decomposition algorithms averaged over $M = 30$ optimization replicas.

	Risk parameters	N	SR	ITER	GAP (%)	Time (s)
SP	$\beta = 0$	10	30	2.0	0.0000	1
		50	30	2.0	0.0000	2
		100	30	2.0	0.0000	3
		500	30	2.0	0.0000	4
		5000	30	2.0	0.0000	43
SP	$\beta = 0.5$, $\alpha = 0.9$	10	30	2.0	0.0000	1
		50	30	2.0	0.0000	2
		100	30	2.0	0.0000	3
		500	30	2.0	0.0000	5
		5000	30	2.0	0.0000	50
SP	$\beta = 1$, $\alpha = 0.9$	10	30	2.1	0.0000	1
		50	30	2.0	0.0000	2
		100	30	2.0	0.0000	3
		500	30	2.0	0.0000	5
		5000	30	2.0	0.0000	51
SP	$\beta = 1$, $\alpha = 0.95$	50	30	2.0	0.0000	2
		100	30	2.0	0.0000	3
		500	30	2.0	0.0000	5
		5000	30	2.0	0.0000	49
SP	$\beta = 1$, $\alpha = 1$	10	30	2.1	0.0000	1
		50	30	2.3	0.0000	2
		100	30	2.7	0.0000	5
		500	30	5.7	0.0000	31
		5000	26	9.3	0.0530	2,556
RO	Γ					
	0	-	1	2	0.0000	1
	10	-	1	3	0.0000	6
	50	-	1	4	0.0000	26
	100	-	1	4	0.0000	47
	150	-	1	4	0.0000	12
	168	-	1	3	0.0000	3

N - sample size used in each optimization replica; SR - number of successful optimization replications that meet the gap stop criterion (out of 30); ITER - average number of iterations of the decomposition algorithm; GAP - gap between the lower bound and upper bound of the decomposition algorithms; Time - wall clock time.

Table S14: Case 2. Week 1. Performance of the decomposition algorithms averaged over $M = 30$ optimization replicas.

	Risk parameters	N	SR	ITER	GAP (%)	Time (s)
SP	$\beta = 0$	10	30	6.3	0.0000	3
		50	30	6.0	0.0000	5
		100	30	6.0	0.0000	9
		500	30	6.0	0.0000	17
		5000	30	6.0	0.0000	180
SP	$\beta = 0.5$, $\alpha = 0.9$	10	30	22.0	0.0000	26
		50	30	6.9	0.0000	10
		100	30	6.2	0.0000	13
		500	30	6.1	0.0000	30
		5000	30	6.1	0.0000	357
SP	$\beta = 1$, $\alpha = 0.9$	10	29	103.2	0.1279	1,396
		50	30	10.7	0.0000	18
		100	30	6.7	0.0000	14
		500	30	6.0	0.0000	26
		5000	30	6.0	0.0000	298
SP	$\beta = 1$, $\alpha = 0.95$	50	30	20.5	0.0000	60
		100	30	9.0	0.0000	22
		500	30	6.0	0.0000	26
		5000	30	6.0	0.0000	294
SP	$\beta = 1$, $\alpha = 1$	10	29	103.2	0.1279	1,399
		50	30	39.8	0.0000	206
		100	30	30.4	0.0000	230
		500	30	18.6	0.0000	461
		5000	28	11.9	0.0263	2,006
RO	Γ					
	0	-	1	2	0.0000	3
	10	-	1	5	0.0000	235
	50	-	0	30	1.1588	10,938
	100	-	0	30	1.2893	10,939
	150	-	0	29	1.3357	10,875
	168	-	0	28	1.2290	11,267

N - sample size used in each optimization replica; SR - number of successful optimization replications that meet the gap stop criterion (out of 30); ITER - average number of iterations of the decomposition algorithm; GAP - gap between the lower bound and upper bound of the decomposition algorithms; Time - wall clock time.

Table S15: Case 1. Week 2. Performance of the decomposition algorithms averaged over $M = 30$ optimization replicas.

	Risk parameters	N	SR	ITER	GAP (%)	Time (s)
SP	$\beta = 0$	10	30	2.0	0.0000	1
		50	30	2.0	0.0000	2
		100	30	2.0	0.0000	3
		500	30	2.0	0.0000	5
		5000	30	2.0	0.0000	44
SP	$\beta = 0.5$, $\alpha = 0.9$	10	30	2.0	0.0000	1
		50	30	2.0	0.0000	2
		100	30	2.0	0.0000	3
		500	30	2.0	0.0000	5
		5000	30	2.0	0.0000	52
SP	$\beta = 1$, $\alpha = 0.9$	10	30	3.8	0.0000	2
		50	30	2.0	0.0000	2
		100	30	2.0	0.0000	3
		500	30	2.0	0.0000	5
		5000	30	2.0	0.0000	51
SP	$\beta = 1$, $\alpha = 0.95$	50	30	2.3	0.0000	2
		100	30	2.0	0.0000	3
		500	30	2.0	0.0000	5
		5000	30	2.0	0.0000	111
SP	$\beta = 1$, $\alpha = 1$	10	30	3.8	0.0000	2
		50	30	3.3	0.0000	3
		100	30	5.8	0.0000	13
		500	30	4.4	0.0000	22
		5000	30	5.5	0.0000	906
RO	Γ					
	0	-	1	2	0.0000	3
	10	-	1	4	0.0000	5
	50	-	1	3	0.0000	10
	100	-	1	4	0.0000	44
	150	-	1	4	0.0000	3
	168	-	1	5	0.0000	4

N - sample size used in each optimization replica; SR - number of successful optimization replications that meet the gap stop criterion (out of 30); ITER - average number of iterations of the decomposition algorithm; GAP - gap between the lower bound and upper bound of the decomposition algorithms; Time - wall clock time.

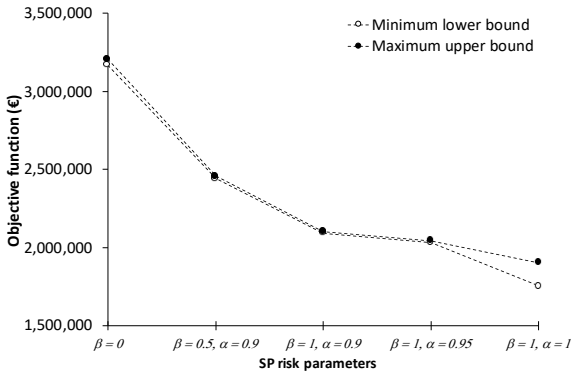
Table S16: Case 2. Week 2. Performance of the decomposition algorithms averaged over $M = 30$ optimization replicas.

	Risk parameters	N	SR	ITER	GAP (%)	Time (s)
SP	$\beta = 0$	10	0	373.9	2.8711	10,849
		50	0	335.6	2.4068	10,838
		100	0	308.1	2.6011	10,851
		500	0	212.3	3.3464	10,849
		5000	0	65.8	5.6553	10,933
SP	$\beta = 0.5$, $\alpha = 0.9$	10	0	75.8	13.0774	11,168
		50	0	50.0	13.3963	11,132
		100	0	40.2	13.1323	11,487
		500	0	27.6	13.2862	11,390
		5000	0	11.0	16.9155	15,062
SP	$\beta = 1$, $\alpha = 0.9$	10	6	53.7	22.3828	9,746
		50	0	32.4	21.8588	11,323
		100	0	28.5	25.3924	11,297
		500	0	23.7	29.2133	11,535
		5000	0	10.4	19.1446	16,210
SP	$\beta = 1$, $\alpha = 0.95$	50	0	30.6	29.9756	11,354
		100	0	28.2	32.9694	11,322
		500	0	23.1	35.9119	11,408
		5000	0	10.5	20.5594	15,658
SP	$\beta = 1$, $\alpha = 1$	10	6	53.6	22.3870	9,741
		50	3	27.9	29.0440	10,376
		100	0	27.1	35.4289	11,364
		500	0	20.4	40.2817	11,587
		5000	0	8.5	37.8558	15,415
RO	Γ					
	0	-	1	2	0.0000	5
	10	-	1	4	0.0000	30
	50	-	0	30	2.4146	11,257
	100	-	0	31	0.1739	11,110
	150	-	1	13	0.0000	1,889
	168	-	1	13	0.0000	1,179

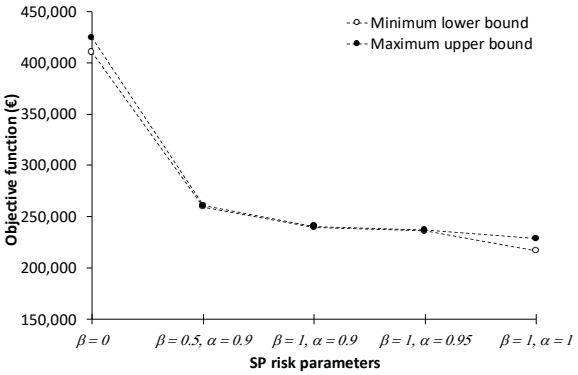
N - sample size used in each optimization replica; SR - number of successful optimization replications that meet the gap stop criterion (out of 30); ITER - average number of iterations of the decomposition algorithm; GAP - gap between the lower bound and upper bound of the decomposition algorithms; Time - wall clock time.

Appendix I. Assessment of the SP solutions

To assess the quality of the SP solutions, we rely on the SAA methodology to provide bounds on the optimal value of the original problem for each set of risk parameters. Specifically, we determine a confidence interval defined by the upper limit of the confidence interval of the estimate of the upper bound on w^* and the lower limit of the confidence interval of the estimate of the lower bound on w^*. Figures S10a and S10b show the confidence intervals for Case 1 and Case 2, respectively, for Week 1. These results indicate that for



(a) Case 1. The relative gaps between the bounds are 1.2%, 0.7%, 0.5%, 0.6%, and 8.4%; from left to right.



(b) Case 2. The relative gaps between the bounds are 3.4%, 0.4% 0.3%, 0.3%, and 5.6%; from left to right.

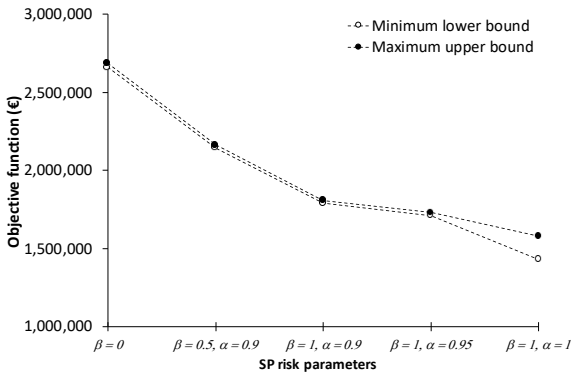
Figure S10: Week 1. Bounds for the true optimal value w^* for each set of risk parameters for a 95% confidence interval. $N = 500$, $M = 30$, $\{T, T'\} = 30$, $N' = 25,000$.

a 95% confidence level, the solutions obtained are close to the true optimal values, which is demonstrated by the relative gaps between the bounds.

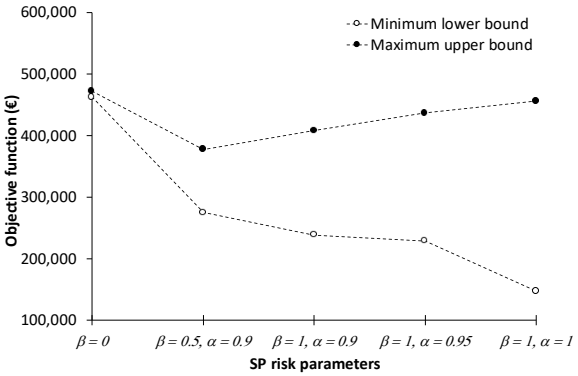
For $\beta = 0$ and $\beta = 1$, $\alpha = 1$, the values of the objective function correspond to the expected profit and CVaR of the profit, whereas the intermediate values of β and α correspond to objective functions with a combination of expected profit and CVaR of the profit. This correspondence explains the trend of the bounds from the left to the right.

For the risk-averse case, $\beta = 1$ and $\alpha = 1$, the gap between the bounds is larger because of two reasons: 1) the gap between the estimates of the lower and upper bounds is larger; and 2) the confidence interval of the estimate of the upper bound is larger than the confidence intervals from the other risk parameters. The larger confidence interval of the estimate of the upper bound is associated with a larger variability on the first-stage solutions obtained in $M = 30$ optimization replications for $\beta = 1$ and $\alpha = 1$, whereas, for example, for $\beta = 0$ the $M = 30$ optimization replications converged to a single solution in Case 1 and 2, for $N = 500$; see the supplementary results in Appendix Appendix D.

The results obtained with Week 2 are presented in Figures S11a and S11b. The results obtained for Case 1 exhibit a similar trend to the ones for Week 1, however, in Case 2 the gap between bounds increased substantially. These results for Case 2 are due to the poor convergence of the L-Shaped method for Case 2, Week 2.



(a) Case 1. The relative gaps between the bounds are 0.9%, 0.9%, 1.0%, 1.0%, and 10.4%; from left to right.



(b) Case 2. The relative gaps between the bounds are 2.0%, 37.3% 71.0%, 90.8%, and 209.7%; from left to right.

Figure S11: Week 2. Bounds for the true optimal value w^* for each set of risk parameters for a 95% confidence interval. $N = 500$, $M = 30$, $\{T, T'\} = 30$, $N' = 25,000$.

Appendix J. Results obtained with the SAA methodology

Appendix J.1. Results from the SP approach with $N = 500$ *and* $M = 30$.

Table S17: Case 1, Week 1. Optimization results for the formulation with $\beta = 0$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
7	3,078,288	2	0.00	4	100	0	0	0	160
11	3,100,959	2	0.00	4	100	0	0	0	160
24	3,117,824	2	0.00	4	100	0	0	0	160
17	3,119,255	2	0.00	4	100	0	0	0	160
4	3,126,409	2	0.00	4	100	0	0	0	160
20	3,131,796	2	0.00	4	100	0	0	0	160
2	3,135,036	2	0.00	4	100	0	0	0	160
16	3,142,733	2	0.00	4	100	0	0	0	160
14	3,147,039	2	0.00	5	100	0	0	0	160
3	3,148,372	2	0.00	5	100	0	0	0	160
28	3,153,705	2	0.00	4	100	0	0	0	160
15	3,170,306	2	0.00	4	100	0	0	0	160
21	3,173,875	2	0.00	4	100	0	0	0	160
8	3,196,010	2	0.00	4	100	0	0	0	160
13	3,196,956	2	0.00	4	100	0	0	0	160
19	3,199,746	2	0.00	4	100	0	0	0	160
30	3,199,782	2	0.00	4	100	0	0	0	160
1	3,201,385	2	0.00	5	100	0	0	0	160
27	3,203,544	2	0.00	4	100	0	0	0	160
23	3,206,883	2	0.00	4	100	0	0	0	160
12	3,209,838	2	0.00	4	100	0	0	0	160
6	3,214,778	2	0.00	5	100	0	0	0	160
10	3,218,465	2	0.00	4	100	0	0	0	160
29	3,223,671	2	0.00	4	100	0	0	0	160
26	3,237,840	2	0.00	4	100	0	0	0	160
5	3,244,176	2	0.00	4	100	0	0	0	160
18	3,255,876	2	0.00	4	100	0	0	0	160
22	3,259,409	2	0.00	4	100	0	0	0	160
9	3,261,767	2	0.00	4	100	0	0	0	160
25	3,274,831	2	0.00	4	100	0	0	0	160

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S18: Case 1, Week 1. Optimization results for the formulation with $\beta = 0.5$, $\alpha = 0.9$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
7	2,410,710	2	0.00	5	100	0	0	155	0
4	2,427,801	2	0.00	5	100	0	0	155	0
17	2,428,278	2	0.00	5	100	0	0	155	0
24	2,428,434	2	0.00	5	100	0	0	155	0
28	2,429,364	2	0.00	5	100	0	0	155	0
2	2,430,276	2	0.00	5	100	0	0	155	0
11	2,430,487	2	0.00	5	100	0	0	155	0
20	2,433,149	2	0.00	5	100	0	0	155	0
19	2,433,720	2	0.00	5	100	0	0	155	0
16	2,433,992	2	0.00	5	100	0	0	155	0
3	2,434,610	2	0.00	5	100	0	0	155	0
15	2,435,269	2	0.00	5	100	0	0	155	0
14	2,437,368	2	0.00	5	100	0	0	155	0
23	2,438,664	2	0.00	5	100	0	0	155	0
13	2,438,807	2	0.00	5	100	0	0	155	0
27	2,440,667	2	0.00	5	100	0	0	155	0
6	2,445,214	2	0.00	5	100	0	0	155	0
12	2,447,029	2	0.00	5	100	0	0	155	0
8	2,449,900	2	0.00	5	100	0	0	155	0
5	2,457,695	2	0.00	5	100	0	0	155	0
30	2,465,297	2	0.00	5	100	0	0	155	0
26	2,467,644	2	0.00	5	100	0	0	155	0
21	2,469,491	2	0.00	5	100	0	0	155	0
29	2,470,247	2	0.00	5	100	0	0	155	0
10	2,471,431	2	0.00	5	100	0	0	155	0
18	2,474,108	2	0.00	5	100	0	0	155	0
9	2,476,679	2	0.00	5	100	0	0	155	0
25	2,482,562	2	0.00	5	100	0	0	155	0
22	2,486,921	2	0.00	5	100	0	0	155	0
1	2,505,719	2	0.00	5	100	0	0	155	0

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S19: Case 1, Week 1. Optimization results for the formulation with $\beta = 1$, $\alpha = 0.9$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
19	2,076,627	2	0.00	5	100	0	0	315	0
23	2,081,075	2	0.00	6	100	0	0	315	0
27	2,082,524	2	0.00	5	100	0	0	315	0
28	2,083,189	2	0.00	5	100	0	0	315	0
7	2,084,137	2	0.00	5	100	0	0	315	0
13	2,084,321	2	0.00	5	100	0	0	315	0
15	2,086,117	2	0.00	5	100	0	0	315	0
12	2,086,933	2	0.00	5	100	0	0	315	0
6	2,087,087	2	0.00	5	100	0	0	315	0
2	2,089,828	2	0.00	5	100	0	0	315	0
4	2,089,874	2	0.00	5	100	0	0	315	0
3	2,090,717	2	0.00	5	100	0	0	315	0
16	2,090,795	2	0.00	5	100	0	0	315	0
5	2,091,541	2	0.00	5	100	0	0	315	0
20	2,092,448	2	0.00	5	100	0	0	315	0
14	2,093,098	2	0.00	5	100	0	0	315	0
24	2,093,149	2	0.00	5	100	0	0	315	0
17	2,093,348	2	0.00	5	100	0	0	315	0
8	2,094,830	2	0.00	5	100	0	0	315	0
11	2,099,166	2	0.00	5	100	0	0	315	0
26	2,100,608	2	0.00	5	100	0	0	315	0
18	2,104,783	2	0.00	5	100	0	0	315	0
25	2,104,902	2	0.00	5	100	0	0	315	0
9	2,105,156	2	0.00	5	100	0	0	315	0
29	2,105,663	2	0.00	5	100	0	0	315	0
10	2,109,383	2	0.00	5	100	0	0	315	0
30	2,109,423	2	0.00	5	100	0	0	315	0
22	2,113,984	2	0.00	5	100	0	0	315	0
21	2,118,535	2	0.00	5	100	0	0	315	0
1	2,150,135	2	0.00	5	100	0	0	260	0

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S20: Case 1, Week 1. Optimization results for the formulation with $\beta = 1$, $\alpha = 0.95$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
23	2,014,825	2	0.00	5	100	0	0	315	0
27	2,020,095	2	0.00	5	100	0	0	315	0
13	2,020,441	2	0.00	5	100	0	0	315	0
19	2,020,864	2	0.00	5	100	0	0	315	0
6	2,024,452	2	0.00	5	100	0	0	315	0
15	2,025,682	2	0.00	5	100	0	0	315	0
7	2,025,891	2	0.00	5	100	0	0	315	0
12	2,026,929	2	0.00	5	100	0	0	315	0
14	2,027,748	2	0.00	5	100	0	0	315	0
28	2,028,724	2	0.00	5	100	0	0	315	0
20	2,030,155	2	0.00	5	100	0	0	315	0
24	2,030,421	2	0.00	5	100	0	0	315	0
4	2,030,831	2	0.00	5	100	0	0	315	0
17	2,031,583	2	0.00	5	100	0	0	315	0
2	2,034,728	2	0.00	5	100	0	0	315	0
16	2,035,310	2	0.00	5	100	0	0	315	0
3	2,036,698	2	0.00	5	100	0	0	315	0
11	2,036,764	2	0.00	5	100	0	0	315	0
5	2,039,686	2	0.00	5	100	0	0	315	0
8	2,040,642	2	0.00	5	100	0	0	315	0
29	2,042,294	2	0.00	5	100	0	0	315	0
26	2,045,748	2	0.00	5	100	0	0	315	0
18	2,047,730	2	0.00	5	100	0	0	315	0
9	2,048,174	2	0.00	5	100	0	0	315	0
25	2,050,981	2	0.00	5	100	0	0	315	0
10	2,054,624	2	0.00	5	100	0	0	315	0
30	2,055,874	2	0.00	5	100	0	0	315	0
22	2,058,026	2	0.00	5	100	0	0	315	0
21	2,067,284	2	0.00	5	100	0	0	315	0
1	2,087,359	2	0.00	5	100	0	0	315	0

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S21: Case 1, Week 1. Optimization results for the formulation with $\beta = 1$, $\alpha = 1$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
12	1,783,267	10	0.00	54	86.9	0	1	315	0
11	1,806,987	2	0.00	5	100	0	0	315	0
26	1,814,875	3	0.00	8	100	0	0	315	0
19	1,824,992	12	0.00	96	92.26	1	1	315	0
15	1,848,261	14	0.00	116	96.43	0	1	315	0
30	1,848,979	15	0.00	148	88.69	0	1	315	0
23	1,849,213	10	0.00	73	91.07	0	1	315	0
4	1,858,044	11	0.00	61	100	0	0	315	0
27	1,870,761	11	0.00	63	95.24	1	1	315	0
3	1,876,680	2	0.00	5	100	0	0	315	0
5	1,878,092	2	0.00	5	100	0	0	315	0
22	1,889,387	6	0.00	22	95.83	0	1	315	0
20	1,890,822	3	0.00	8	100	0	0	315	0
9	1,892,352	2	0.00	5	100	0	0	315	0
28	1,893,542	14	0.00	95	92.26	1	1	315	0
25	1,894,784	8	0.00	34	100	0	0	315	0
13	1,898,621	7	0.00	29	98.21	0	1	315	0
24	1,903,589	3	0.00	8	100	0	0	315	0
6	1,906,827	2	0.00	5	100	0	0	315	0
10	1,910,613	2	0.00	5	100	0	0	315	0
8	1,910,617	2	0.00	5	100	0	0	315	0
16	1,911,058	3	0.00	8	100	0	0	315	0
21	1,914,069	3	0.00	8	100	0	0	315	0
2	1,918,108	2	0.00	5	100	0	0	315	0
17	1,918,410	6	0.00	22	100	0	0	315	0
7	1,920,187	4	0.00	13	100	0	0	315	0
29	1,922,000	3	0.00	8	100	0	0	315	0
18	1,934,331	3	0.00	8	100	0	0	315	0
14	1,937,937	3	0.00	8	100	0	0	315	0
1	1,968,190	2	0.00	5	100	0	0	315	0

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S22: Case 2, Week 1. Optimization results for the formulation with $\beta = 0$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
7	379,038	6	0.00	17	100	1	0	0	160
11	386,098	6	0.00	17	100	1	0	0	160
24	393,393	6	0.00	17	100	1	0	0	160
17	394,248	6	0.00	17	100	1	0	0	160
4	396,543	6	0.00	17	100	1	0	0	160
20	397,845	6	0.00	17	100	1	0	0	160
2	398,937	6	0.00	17	100	1	0	0	160
16	402,244	6	0.00	17	100	1	0	0	160
14	402,376	6	0.00	17	100	1	0	0	160
3	404,888	6	0.00	16	100	1	0	0	160
28	405,871	6	0.00	17	100	1	0	0	160
15	412,871	6	0.00	17	100	1	0	0	160
21	414,024	6	0.00	17	100	1	0	0	160
8	421,739	6	0.00	17	100	1	0	0	160
13	422,243	6	0.00	17	100	1	0	0	160
1	422,718	7	0.00	20	100	1	0	0	160
30	423,331	6	0.00	17	100	1	0	0	160
19	424,465	6	0.00	17	100	1	0	0	160
27	424,567	6	0.00	17	100	1	0	0	160
23	425,526	6	0.00	17	100	1	0	0	160
12	425,587	6	0.00	17	100	1	0	0	160
6	427,701	6	0.00	17	100	1	0	0	160
10	429,332	6	0.00	17	100	1	0	0	160
29	431,214	6	0.00	17	100	1	0	0	160
26	435,081	6	0.00	16	100	1	0	0	160
5	438,943	6	0.00	17	100	1	0	0	160
18	441,966	6	0.00	17	100	1	0	0	160
22	445,356	6	0.00	17	100	1	0	0	160
9	445,841	6	0.00	17	100	1	0	0	160
25	450,188	6	0.00	17	100	1	0	0	160

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S23: Case 2, Week 1. Optimization results for the formulation with $\beta = 0.5$, $\alpha = 0.9$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
20	257,748	6	0.00	31	100	1	0	100	49.67
14	257,981	6	0.00	30	100	1	0	100	47.57
16	258,025	6	0.00	31	100	1	0	100	46.9
11	258,245	6	0.00	30	100	1	0	100	47.78
12	258,349	6	0.00	30	100	1	0	100	49.43
28	258,411	6	0.00	30	100	1	0	100	48.6
7	258,778	7	0.00	39	100	1	0	100	45.11
2	258,898	6	0.00	30	100	1	0	100	48.54
26	259,360	6	0.00	30	100	1	0	100	55
24	259,594	6	0.00	30	100	1	0	100	51.1
4	259,682	6	0.00	30	100	1	0	100	48.31
17	259,779	6	0.00	29	100	1	0	100	49.86
29	259,965	6	0.00	30	100	1	0	100	53.95
15	260,170	6	0.00	31	100	1	0	100	50.39
10	260,267	6	0.00	29	100	1	0	100	55
3	260,293	6	0.00	29	100	1	0	100	49.55
27	260,300	6	0.00	29	100	1	0	100	48.91
23	260,378	6	0.00	30	100	1	0	100	52
1	260,427	7	0.00	36	100	1	0	100	55
25	260,772	6	0.00	30	100	1	0	100	55
21	260,784	6	0.00	28	100	1	0	100	54.36
6	260,793	6	0.00	30	100	1	0	100	52.16
19	260,803	6	0.00	31	100	1	0	100	49.61
8	260,918	6	0.00	30	100	1	0	100	54.74
5	261,027	6	0.00	30	100	1	0	100	53.54
18	261,690	6	0.00	29	100	1	0	100	55
13	261,790	6	0.00	29	100	1	0	100	50.38
30	261,802	6	0.00	30	100	1	0	100	55
9	262,194	6	0.00	29	100	1	0	100	55
22	263,220	6	0.00	29	100	1	0	100	55

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S24: Case 2, Week 1. Optimization results for the formulation with $\beta = 1$, $\alpha = 0.9$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
20	238,248	6	0.00	26	100	1	0	100	38.27
12	238,967	6	0.00	26	100	1	0	100	38.58
10	239,050	6	0.00	25	100	1	0	100	38.48
29	239,153	6	0.00	25	100	1	0	100	38.32
30	239,174	6	0.00	25	100	1	0	100	39.59
24	239,257	6	0.00	26	100	1	0	100	39.65
14	239,289	6	0.00	26	100	1	0	100	38.59
28	239,433	6	0.00	26	100	1	0	100	37.52
16	239,469	6	0.00	26	100	1	0	100	37.98
11	239,732	6	0.00	26	100	1	0	100	39.15
17	239,767	6	0.00	25	100	1	0	100	39.15
27	239,796	6	0.00	25	100	1	0	100	38.8
9	239,806	6	0.00	25	100	1	0	100	38.76
8	239,847	6	0.00	26	100	1	0	100	39.13
23	239,886	6	0.00	26	100	1	0	100	37.97
3	239,897	6	0.00	26	100	1	0	100	38.67
25	239,928	6	0.00	29	100	1	0	100	38.06
21	240,030	6	0.00	26	100	1	0	100	38.64
26	240,099	6	0.00	26	100	1	0	100	38.67
19	240,265	6	0.00	26	100	1	0	100	37.53
1	240,586	7	0.00	33	100	1	0	100	38.58
5	240,662	6	0.00	26	100	1	0	100	38.72
4	240,815	6	0.00	26	100	1	0	100	37.84
2	240,869	6	0.00	25	100	1	0	100	38.52
7	240,941	6	0.00	26	100	1	0	100	39.18
15	240,967	6	0.00	25	100	1	0	100	38.09
18	241,011	6	0.00	26	100	1	0	100	39.85
22	241,191	6	0.00	26	100	1	0	100	39.25
6	241,577	6	0.00	26	100	1	0	100	39.17
13	242,038	6	0.00	25	100	1	0	100	38.82

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S25: Case 2, Week 1. Optimization results for the formulation with $\beta = 1$, $\alpha = 0.95$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
20	234,465	6	0.00	26	100	1	0	100	38.44
30	234,721	6	0.00	25	100	1	0	100	39.28
10	234,727	6	0.00	26	100	1	0	100	38.92
12	234,846	6	0.00	25	100	1	0	100	37.74
24	235,168	6	0.00	25	100	1	0	100	39.71
8	235,489	6	0.00	26	100	1	0	100	39.11
21	235,548	6	0.00	26	100	1	0	100	38.17
14	235,640	6	0.00	25	100	1	0	100	37.47
25	235,774	6	0.00	25	100	1	0	100	37.36
27	235,791	6	0.00	25	100	1	0	100	38.4
29	235,932	6	0.00	26	100	1	0	100	38.22
3	235,974	6	0.00	25	100	1	0	100	37.48
23	236,122	6	0.00	25	100	1	0	100	37.04
28	236,150	6	0.00	25	100	1	0	100	36.33
11	236,219	6	0.00	25	100	1	0	100	38.89
26	236,223	6	0.00	25	100	1	0	100	37.58
17	236,303	6	0.00	24	100	1	0	100	38.27
16	236,338	6	0.00	26	100	1	0	100	36.19
19	236,426	6	0.00	25	100	1	0	100	38.3
5	236,484	6	0.00	26	100	1	0	100	39.47
9	236,626	6	0.00	25	100	1	0	100	37.39
7	237,146	6	0.00	24	100	1	0	100	39.49
18	237,207	6	0.00	25	100	1	0	100	39.38
22	237,237	6	0.00	26	100	1	0	100	38.39
15	237,293	6	0.00	26	100	1	0	100	37.45
2	237,322	6	0.00	25	100	1	0	100	38.52
1	237,420	7	0.00	30	100	1	0	100	37.8
6	237,454	6	0.00	26	100	1	0	100	39.42
4	237,530	6	0.00	25	100	1	0	100	37.38
13	237,836	6	0.00	25	100	1	0	100	38.02

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S26: Case 2, Week 1. Optimization results for the formulation with $\beta = 1$, $\alpha = 1$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
12	218,757	6	0.00	24	100	1	0	100	43.36
30	220,833	9	0.00	47	100	1	0	100	37.83
20	222,004	6	0.00	24	100	1	0	100	34.84
24	223,950	19	0.00	199	100	1	0	100	36.27
10	224,680	26	0.00	407	100	1	0	100	44.26
13	225,354	15	0.00	129	100	1	0	100	44.65
14	225,985	15	0.00	133	100	1	0	100	33.17
8	226,144	24	0.00	332	100	1	0	100	43.19
6	226,291	6	0.00	25	100	1	0	100	40.97
5	226,837	7	0.00	32	100	1	0	100	38.08
9	227,104	12	0.00	80	100	1	0	100	36.91
29	227,114	19	0.00	215	100	1	0	100	41.73
27	227,267	37	0.00	819	100	1	0	100	36.07
21	227,566	6	0.00	24	100	1	0	100	36.37
25	228,217	36	0.00	943	100	1	0	100	37.87
4	228,328	47	0.00	1,826	100	1	0	100	38.04
7	228,755	6	0.00	24	100	1	0	100	38.58
22	228,859	74	0.00	6,095	100	1	0	100	34.6
16	228,926	6	0.00	25	100	1	0	100	36.9
26	229,088	32	0.00	685	100	1	0	100	36.73
28	229,236	9	0.00	53	100	1	0	100	33.84
19	229,549	32	0.00	652	100	1	0	100	36.78
1	229,644	20	0.00	231	100	1	0	100	35.86
3	229,761	13	0.00	104	100	1	0	100	37.3
17	230,121	9	0.00	49	100	1	0	100	41.82
15	230,258	19	0.00	244	100	1	0	100	34.53
11	230,308	19	0.00	198	100	1	0	100	39.65
2	230,839	6	0.00	26	100	1	0	100	40.5
18	231,640	18	0.00	173	100	1	0	100	37.02
23	232,142	6	0.00	24	100	1	0	100	36.72

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S27: Case 1, Week 2. Optimization results for the formulation with $\beta = 0$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
7	2,604,233	2	0.00	4	100	0	0	0	160
17	2,620,361	2	0.00	4	100	0	0	0	160
11	2,626,283	2	0.00	4	100	0	0	0	160
24	2,631,787	2	0.00	4	100	0	0	0	160
3	2,632,211	2	0.00	4	100	0	0	0	160
2	2,643,010	2	0.00	4	100	0	0	0	160
16	2,647,354	2	0.00	4	100	0	0	0	160
15	2,648,129	2	0.00	4	100	0	0	0	160
21	2,648,863	2	0.00	4	100	0	0	0	160
28	2,650,303	2	0.00	7	100	0	0	0	160
20	2,653,771	2	0.00	4	100	0	0	0	160
19	2,657,708	2	0.00	4	100	0	0	0	160
4	2,661,912	2	0.00	4	100	0	0	0	160
14	2,664,349	2	0.00	4	100	0	0	0	160
30	2,669,870	2	0.00	4	100	0	0	0	160
13	2,671,769	2	0.00	4	100	0	0	0	160
27	2,672,568	2	0.00	4	100	0	0	0	160
8	2,681,209	2	0.00	4	100	0	0	0	160
12	2,685,021	2	0.00	4	100	0	0	0	160
23	2,686,597	2	0.00	4	100	0	0	0	160
6	2,690,264	2	0.00	4	100	0	0	0	160
5	2,696,082	2	0.00	4	100	0	0	0	160
29	2,702,453	2	0.00	4	100	0	0	0	160
1	2,705,720	3	0.00	8	100	0	0	0	160
9	2,711,809	2	0.00	4	100	0	0	0	160
18	2,714,041	2	0.00	4	100	0	0	0	160
10	2,714,877	2	0.00	4	100	0	0	0	160
26	2,717,774	2	0.00	4	100	0	0	0	160
25	2,721,459	2	0.00	4	100	0	0	0	160
22	2,725,259	2	0.00	4	100	0	0	0	160

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S28: Case 1, Week 2. Optimization results for the formulation with $\beta = 0.5$, $\alpha = 0.9$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
3	2,117,817	2	0.00	5	100	0	0	50	55
24	2,119,738	2	0.00	5	100	0	0	50	55
7	2,119,850	2	0.00	5	100	0	0	50	55
19	2,122,621	2	0.00	5	100	0	0	50	55
17	2,123,902	2	0.00	5	100	0	0	50	55
11	2,127,937	2	0.00	5	100	0	0	50	55
6	2,131,610	2	0.00	5	100	0	0	50	55
28	2,133,940	2	0.00	5	100	0	0	50	55
2	2,137,153	2	0.00	5	100	0	0	50	55
16	2,137,512	2	0.00	5	100	0	0	50	55
5	2,140,159	2	0.00	5	100	0	0	50	55
15	2,140,742	2	0.00	5	100	0	0	50	55
20	2,143,321	2	0.00	5	100	0	0	50	55
4	2,145,341	2	0.00	5	100	0	0	50	55
21	2,145,784	2	0.00	5	100	0	0	50	55
13	2,145,998	2	0.00	5	100	0	0	50	55
27	2,146,332	2	0.00	5	100	0	0	50	55
23	2,149,386	2	0.00	5	100	0	0	50	55
9	2,155,694	2	0.00	5	100	0	0	50	55
30	2,160,677	2	0.00	5	100	0	0	50	55
8	2,160,755	2	0.00	5	100	0	0	50	55
18	2,163,509	2	0.00	6	100	0	0	50	55
14	2,165,378	2	0.00	5	100	0	0	50	55
12	2,168,133	2	0.00	5	100	0	0	50	55
25	2,186,320	2	0.00	5	100	0	0	50	55
26	2,187,711	2	0.00	5	100	0	0	50	55
22	2,189,910	2	0.00	5	100	0	0	50	55
29	2,193,321	2	0.00	5	100	0	0	50	55
10	2,201,951	2	0.00	5	100	0	0	50	55
1	2,216,005	3	0.00	9	100	0	0	50	55

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S29: Case 1, Week 2. Optimization results for the formulation with $\beta = 1$, $\alpha = 0.9$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
19	1,765,064	2	0.00	5	100	0	0	155	0
6	1,765,595	2	0.00	5	100	0	0	210	0
5	1,769,182	2	0.00	5	100	0	0	155	0
3	1,773,377	2	0.00	5	100	0	0	155	0
24	1,773,951	2	0.00	5	100	0	0	155	0
9	1,781,951	2	0.00	5	100	0	0	155	0
28	1,785,166	2	0.00	5	100	0	0	155	0
23	1,785,309	2	0.00	5	100	0	0	155	0
7	1,786,598	2	0.00	5	100	0	0	155	0
27	1,787,206	2	0.00	5	100	0	0	155	0
17	1,788,645	2	0.00	5	100	0	0	155	0
18	1,789,648	2	0.00	5	100	0	0	155	0
16	1,789,841	2	0.00	5	100	0	0	155	0
13	1,790,073	2	0.00	5	100	0	0	155	0
11	1,790,082	2	0.00	5	100	0	0	155	0
2	1,791,606	2	0.00	5	100	0	0	155	0
4	1,792,455	2	0.00	5	100	0	0	155	0
20	1,794,615	2	0.00	5	100	0	0	155	0
15	1,795,002	2	0.00	5	100	0	0	155	0
21	1,799,752	2	0.00	5	100	0	0	155	0
8	1,806,422	2	0.00	5	100	0	0	155	0
30	1,808,593	2	0.00	5	100	0	0	155	0
12	1,814,114	2	0.00	5	100	0	0	155	0
25	1,814,197	2	0.00	5	100	0	0	155	0
14	1,818,975	2	0.00	5	100	0	0	155	0
22	1,819,699	2	0.00	5	100	0	0	155	0
26	1,823,726	2	0.00	5	100	0	0	155	0
29	1,837,284	2	0.00	5	100	0	0	155	0
10	1,843,654	2	0.00	5	100	0	0	155	0
1	1,865,242	3	0.00	8	100	0	0	155	0

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S30: Case 1, Week 2. Optimization results for the formulation with $\beta = 1$, $\alpha = 0.95$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
5	1,683,761	2	0.00	5	100	0	0	260	0
6	1,687,531	2	0.00	5	100	0	0	260	0
23	1,692,596	2	0.00	5	100	0	0	260	0
19	1,699,717	2	0.00	5	100	0	0	260	0
24	1,700,288	2	0.00	5	100	0	0	260	0
27	1,701,354	2	0.00	5	100	0	0	260	0
9	1,701,612	2	0.00	5	100	0	0	260	0
28	1,705,078	2	0.00	5	100	0	0	260	0
16	1,707,270	2	0.00	5	100	0	0	260	0
7	1,711,544	2	0.00	5	100	0	0	210	0
18	1,713,376	2	0.00	5	100	0	0	260	0
13	1,714,584	2	0.00	5	100	0	0	260	0
3	1,715,365	2	0.00	5	100	0	0	260	0
20	1,715,568	2	0.00	5	100	0	0	260	0
11	1,716,144	2	0.00	5	100	0	0	260	0
30	1,716,702	2	0.00	5	100	0	0	260	0
2	1,717,297	2	0.00	5	100	0	0	210	0
4	1,719,527	2	0.00	5	100	0	0	210	0
15	1,721,245	2	0.00	5	100	0	0	260	0
17	1,722,128	2	0.00	5	100	0	0	260	0
8	1,727,201	2	0.00	5	100	0	0	260	0
25	1,728,304	2	0.00	5	100	0	0	210	0
21	1,730,942	2	0.00	5	100	0	0	210	0
26	1,739,322	2	0.00	5	100	0	0	210	0
14	1,742,210	2	0.00	5	100	0	0	210	0
12	1,742,756	2	0.00	5	100	0	0	210	0
22	1,744,431	2	0.00	5	100	0	0	210	0
29	1,748,966	2	0.00	5	100	0	0	210	0
10	1,756,191	2	0.00	5	100	0	0	210	0
1	1,794,332	3	0.00	9	100	0	0	155	0

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S31: Case 1, Week 2. Optimization results for the formulation with $\beta = 1$, $\alpha = 1$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
19	1,454,470	4	0.00	21	100	0	0	315	0
25	1,510,987	11	0.00	62	95.24	1	1	315	0
28	1,519,704	3	0.00	9	100	0	0	315	0
17	1,520,896	12	0.00	100	100	0	0	315	0
27	1,522,801	7	0.00	32	100	0	0	315	0
5	1,524,831	2	0.00	8	100	0	0	315	0
13	1,527,372	2	0.00	6	100	0	0	315	0
7	1,528,090	3	0.00	10	100	0	0	315	0
30	1,528,516	5	0.00	18	100	0	0	315	0
3	1,537,998	2	0.00	9	100	0	0	315	0
22	1,543,869	7	0.00	30	100	0	0	315	0
8	1,544,393	2	0.00	5	100	0	0	315	0
6	1,547,296	9	0.00	50	100	0	0	315	0
12	1,548,427	13	0.00	102	100	0	0	315	0
24	1,561,759	2	0.00	6	100	0	0	260	0
4	1,565,939	2	0.00	7	100	0	0	315	0
23	1,566,855	2	0.00	6	100	0	0	315	0
26	1,568,833	6	0.00	30	100	0	0	260	0
9	1,581,039	4	0.00	16	100	0	0	260	0
16	1,581,480	3	0.00	10	100	0	0	260	0
15	1,587,210	4	0.00	13	100	0	0	260	0
2	1,589,208	2	0.00	6	100	0	0	260	0
11	1,597,446	2	0.00	5	100	0	0	260	0
20	1,601,428	2	0.00	6	100	0	0	294.72	0
29	1,602,649	9	0.00	49	100	0	0	260	0
14	1,609,729	2	0.00	7	100	0	0	260	0
21	1,610,322	2	0.00	13	100	0	0	260	0
18	1,619,745	2	0.00	5	100	0	0	260	0
1	1,632,789	3	0.00	20	100	0	0	260	0
10	1,656,466	2	0.00	6	100	0	0	260	0

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S32: Case 2, Week 2. Optimization results for the formulation with $\beta = 0$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
7	440,469	204	3.88	10,821	93.45	6	5	0	160
17	447,766	211	3.70	10,887	93.45	6	5	0	160
3	451,901	175	3.91	10,900	94.64	6	5	0	160
11	452,428	209	3.65	10,814	93.45	6	5	0	160
24	455,163	208	3.50	10,888	93.45	6	5	0	160
15	457,008	210	3.58	10,822	93.45	6	5	0	160
21	457,148	214	3.52	10,830	93.45	6	5	0	160
16	457,890	214	3.29	10,892	93.45	6	5	0	160
2	458,906	205	2.67	10,852	97.02	3	2	0	160
28	461,131	213	3.48	10,812	93.45	6	5	0	160
19	461,467	215	3.42	10,837	93.45	6	5	0	160
20	461,631	210	3.59	10,819	93.45	6	5	0	160
14	464,045	213	3.49	10,806	93.45	6	5	0	160
4	465,105	195	3.59	10,845	93.45	6	5	0	160
13	465,180	217	3.34	10,837	93.45	6	5	0	160
27	466,933	218	3.33	10,888	93.45	6	5	0	160
30	467,315	216	3.35	10,827	93.45	6	5	0	160
8	471,506	211	3.44	10,851	93.45	6	5	0	160
12	473,450	218	3.14	10,839	93.45	6	5	0	160
23	473,626	218	3.10	10,825	93.45	6	5	0	160
5	473,841	218	3.14	10,850	93.45	6	5	0	160
6	475,196	217	3.26	10,889	93.45	6	5	0	160
1	477,630	221	2.73	10,894	88.69	5	4	0	160
29	479,837	217	3.24	10,871	93.45	6	5	0	160
9	482,099	218	3.28	10,849	93.45	6	5	0	160
18	482,565	219	3.22	10,805	93.45	6	5	0	160
26	484,549	217	2.94	10,890	93.45	6	5	0	160
25	484,963	217	3.25	10,819	93.45	6	5	0	160
10	487,219	215	3.17	10,861	93.45	6	5	0	160
22	488,866	215	3.19	10,841	93.45	6	5	0	160

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S33: Case 2, Week 2. Optimization results for the formulation with $\beta = 0.5$, $\alpha = 0.9$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
1	345,172	35	12.35	11,615	83.33	10	9	57.23	0
7	367,665	30	13.10	11,413	87.5	13	12	61.11	0
13	367,704	31	12.55	11,181	80.36	24	23	60.03	0
12	370,471	30	13.20	10,831	85.12	16	15	61.15	0
30	371,740	30	13.14	11,094	83.33	17	16	61.05	0
15	372,266	21	13.01	11,389	79.76	22	21	60.59	0
21	372,748	29	12.83	11,128	85.71	14	13	61.56	0
25	372,900	34	14.55	10,863	89.29	10	9	62.13	0
2	372,907	28	13.35	11,487	85.71	15	14	61.59	0
9	373,356	29	13.23	11,255	88.1	11	10	62.02	0
16	373,650	28	13.26	11,235	86.31	14	13	61.78	0
3	373,857	26	13.49	10,940	88.1	17	16	62.11	0
29	374,809	25	12.78	12,359	85.71	16	15	61.87	0
6	374,869	27	13.24	11,558	87.5	12	11	62.14	0
24	374,997	25	13.66	11,731	86.31	14	13	61.98	0
27	375,017	31	13.65	11,167	87.5	12	11	62.16	0
23	376,148	30	13.22	11,130	88.69	11	10	62.51	0
4	376,921	27	13.93	11,245	84.52	16	15	61.99	0
18	376,963	21	12.70	11,531	82.14	14	13	61.61	0
19	377,155	25	13.72	12,014	85.71	13	12	62.19	0
8	377,445	24	13.23	10,952	85.12	14	13	62.15	0
14	378,806	30	13.72	11,388	87.5	13	12	62.71	0
5	378,881	25	12.94	11,292	80.36	22	21	61.62	0
11	379,064	23	12.93	11,876	87.5	11	10	62.74	0
26	379,120	23	13.32	11,927	86.31	13	12	62.56	0
17	379,123	27	12.36	11,479	85.12	15	14	62.39	0
28	379,248	31	13.93	11,334	86.31	16	15	62.59	0
20	382,229	24	13.07	11,985	86.31	18	17	63.03	0
10	383,043	25	13.06	10,983	88.69	13	12	63.5	0
22	383,885	33	15.08	11,318	87.5	12	11	63.43	0

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S34: Case 2, Week 2. Optimization results for the formulation with $\beta = 1$, $\alpha = 0.9$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
1	297,409	29	11.81	11,086	80.95	10	9	50	0
12	354,620	24	20.86	11,085	82.74	17	16	58.51	0
14	366,625	28	25.25	11,157	87.5	14	13	60.97	0
26	366,778	29	23.27	11,349	81.55	19	18	60.08	0
8	374,727	26	23.63	10,907	86.9	11	10	62.03	0
10	375,532	18	23.30	11,969	83.33	17	16	61.6	0
28	380,086	20	26.55	11,970	86.9	14	13	62.8	0
17	382,957	23	27.03	11,714	85.12	18	17	62.94	0
23	383,571	27	25.80	11,033	85.12	16	15	63.03	0
22	385,449	30	26.86	11,149	84.52	15	14	63.2	0
18	385,910	31	27.59	11,891	85.71	16	15	63.46	0
29	387,198	25	27.02	11,125	85.12	14	13	63.54	0
2	387,878	27	28.10	10,873	85.71	16	15	63.74	0
15	397,085	20	29.29	11,963	86.31	15	14	65.15	0
6	402,478	24	29.64	10,926	82.74	12	11	65.35	0
21	406,058	21	31.28	11,023	79.76	20	19	65.42	0
3	406,490	19	31.46	12,560	87.5	14	13	66.69	0
11	408,433	24	31.25	11,647	85.71	16	15	66.7	0
16	411,015	28	33.22	11,034	84.52	18	17	66.88	0
13	412,305	23	31.50	10,999	81.55	16	15	66.59	0
27	413,760	21	32.68	11,497	87.5	14	13	67.73	0
20	415,762	21	32.51	11,793	85.71	15	14	67.73	0
24	416,083	19	32.31	10,924	86.31	15	14	67.87	0
30	416,444	21	34.67	14,381	88.69	12	11	68.29	0
9	422,391	19	33.96	11,442	86.9	14	13	68.87	0
4	423,943	21	34.19	13,142	83.33	16	15	68.54	0
25	425,058	23	35.32	11,147	85.12	16	15	68.97	0
19	427,808	21	34.43	11,185	86.9	17	16	69.65	0
5	436,887	28	35.55	11,536	83.33	18	17	70.39	0
7	439,296	20	36.06	11,555	86.31	13	12	71.2	0

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S35: Case 2, Week 2. Optimization results for the formulation with $\beta = 1$, $\alpha = 0.95$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
1	361,069	31	30.39	10,887	80.95	8	7	59.12	0
10	370,700	19	23.88	12,257	82.14	14	13	60.71	0
26	393,178	18	29.06	12,224	82.74	22	21	64.05	0
8	403,750	20	32.17	11,512	83.33	21	20	65.66	0
23	404,095	20	32.66	12,250	86.9	17	16	66.25	0
30	407,101	29	33.36	10,818	83.33	14	13	66.12	0
27	410,304	27	34.90	10,966	83.93	14	13	66.67	0
3	411,401	30	33.14	11,194	85.71	14	13	67.11	0
17	412,959	21	32.37	11,457	79.17	17	16	66.31	0
2	413,221	28	34.36	10,882	82.14	20	19	66.83	0
18	414,767	27	34.37	11,521	83.33	18	17	67.23	0
28	416,113	23	35.43	11,839	84.52	17	16	67.61	0
15	420,685	17	34.09	11,004	80.95	18	17	67.69	0
11	421,465	26	33.57	11,185	83.93	16	15	68.27	0
20	424,252	21	34.71	11,287	85.12	16	15	68.86	0
21	426,683	29	35.59	11,301	82.14	15	14	68.74	0
22	429,372	29	38.08	11,209	85.71	16	15	69.69	0
4	430,262	19	36.63	11,258	86.31	13	12	69.9	0
7	431,025	17	38.21	11,162	82.14	20	19	69.38	0
29	431,517	15	35.59	13,246	79.76	25	24	69.1	0
16	438,651	20	39.58	11,392	84.52	17	16	70.84	0
12	443,088	26	38.93	11,489	88.69	11	10	72.1	0
14	444,535	26	38.87	10,949	87.5	16	15	72.14	0
19	446,983	22	37.94	11,190	83.93	13	12	71.92	0
5	453,358	24	40.26	10,829	83.33	13	12	72.74	0
9	455,013	23	39.36	11,722	87.5	15	14	73.64	0
6	455,990	28	40.78	11,030	88.69	10	9	73.95	0
13	466,253	22	41.84	11,130	86.9	18	17	75.17	0
24	472,608	13	42.03	12,084	83.93	22	21	75.63	0
25	481,095	22	45.18	10,954	82.74	19	18	76.65	0

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.

Table S36: Case 2, Week 2. Optimization results for the formulation with $\beta = 1$, $\alpha = 1$. The results are ordered by the value of the objective function. $M = 30$, $N = 500$.

					First-stage variables (aggregated)				
m	$\hat{w}_{N,m}$ (€)	ITER	GAP (%)	T (s)	UT (%)	SUP	SD	SELLC (MW)	BUYC (MW)
12	341,646	21	25.84	11,405	73.81	22	21	55.26	0
26	353,082	26	26.08	11,380	71.43	22	21	56.53	0
8	355,955	22	25.81	10,985	75	18	18	57.49	0
28	407,904	18	38.64	11,533	71.43	20	20	64.37	0
29	412,031	14	34.82	13,690	76.79	14	13	65.79	0
17	415,133	25	33.47	11,235	76.19	20	19	66.15	0
20	416,103	23	37.51	11,243	76.79	19	18	66.38	0
4	420,918	18	37.16	11,080	79.76	15	14	67.54	0
18	421,309	24	37.22	11,211	78.57	17	16	67.41	0
23	427,581	21	38.22	11,408	74.4	15	14	67.65	0
15	434,127	18	38.10	11,620	76.79	16	15	68.97	0
1	443,200	39	48.86	11,158	76.19	8	7	70.14	0
10	443,821	20	40.27	11,569	76.19	13	12	70.24	0
13	446,807	18	43.56	10,853	72.02	14	13	70.02	0
11	447,129	24	41.10	11,377	77.98	17	17	71.02	0
25	448,386	14	44.20	12,157	72.02	21	20	70.27	0
27	448,950	16	40.65	11,337	75.6	21	20	70.92	0
9	454,789	13	41.73	16,898	79.17	19	18	72.32	0
24	455,693	24	41.88	11,066	77.38	16	15	72.15	0
30	457,360	12	41.81	11,967	73.21	25	24	71.77	0
6	458,802	23	41.69	11,364	76.19	12	11	72.38	0
19	462,397	17	43.17	11,760	79.17	17	16	73.39	0
7	468,938	21	44.76	10,940	80.36	15	15	74.51	0
3	470,288	19	43.05	10,843	74.4	25	24	73.81	0
2	470,688	17	43.16	11,752	81.55	21	20	74.97	0
22	473,381	18	48.31	10,989	75.6	16	15	74.4	0
14	477,908	23	44.18	11,498	83.93	13	12	76.36	0
5	495,176	23	48.67	11,315	82.14	16	15	78.56	0
21	497,313	18	46.75	11,139	76.19	21	20	77.95	0
16	506,892	22	47.79	10,852	80.95	17	16	80.06	0

m - Optimization replication number, $\hat{w}_{N,m}$ - objective function value for optimization replication m, ITER - number of iterations of the L-Shaped method, GAP - gap between the upper and lower bound withing the L-Shaped method, T - elapsed wall-clock time, UT - percentage of up-time of the thermal unit, SUP - number of startups of the thermal unit, SD - number of shutdowns of the thermal unit, SELLC - power sold through contract, BUYC - power bought through contract.